

# Engineering heme proteins for olefin and carbon–hydrogen bond functionalization reactions

Thesis by  
Ruijie (Kelly) Zhang

In Partial Fulfillment of the Requirements for the  
degree of  
Doctor of Philosophy

The Caltech logo, featuring the word "Caltech" in a bold, orange, sans-serif font.

CALIFORNIA INSTITUTE OF TECHNOLOGY  
Pasadena, California

2019  
Defended February 22, 2019

© 2019

Ruijie (Kelly) Zhang  
ORCID: 0000-0002-7251-5527

## ACKNOWLEDGEMENTS

To my advisor and role model, Prof. Frances Arnold, who taught me to be fearless in the pursuit of my ideas. Her approach to science, in which results are scrutinized both for accuracy and broader implications, challenges me to continue to establish *new concepts* as I progress in my scientific career. Frances has a skill for asking tough questions, then helping her students find their own path in exploring the ideas.

Thank you to my committee members, Prof. Greg Fu, Prof. Shu-ou Shan, and Prof. Brian Stoltz. Each member brings his/her own expertise, which they have enthusiastically shared. In particular, Greg, with his impressive knowledge of chemistry and high attention to detail, has provided me with invaluable thought-provoking comments throughout my graduate school career. Brian has called attention to critical details in my research and asked questions to help me understand their importance. Shu-ou's understanding of enzyme mechanisms is something I greatly admire and which I have only just scratched the surface of with her help. Looking back, I should have visited my committee members more frequently in order to take full advantage of their combined knowledge.

The success of a scientist is significantly impacted by the quality of her peers. I have had the privilege to work with and learn from many outstanding scientists at Caltech, especially Dr. Sabine Brinkmann-Chen, Dr. Chris Farwell, Dr. John McIntosh, Dr. Chris Prier, Dr. David Romney, Dr. Andrew Buller, Dr. Hans Renata, Lena Wohlschlager, Kai Chen, and Dr. Xiongyi Huang. I have many great memories from the third (and second) floor of Spalding labs. The Arnold group is a vibrant and dynamic scientific community; people are always interested in discussion and willing to challenge others to be better versions of their scientific selves ("well actually..."). Experts including Dr. Scott Virgil and Dr. David VanderVelde have been generous with their time and patient in sharing their expertise. Mentors, friends, and colleagues contribute a large role to the scientist I am today.

To my friends from Chicago, particularly Mehnaaz Chowdhury and Auri Pope (and Sam Ho), who made these years in Los Angeles unforgettable. And finally, to my family, Shurong Yang, Jianhua Zhang, Mellissa, and Richard, for their endless love and support.

## ABSTRACT

One of the most important challenges in chemistry is the creation of new catalysts. Nature excels at this: constructed from biologically available elements, enzymes are versatile catalysts which adapt quickly to changing environments in order to sustain life. The combination of adaptable proteins with abiological reagents from synthetic chemistry affords a new direction for catalyst development. This thesis describes new enzymes, derived from a cytochrome P450 monooxygenase, which catalyze nitrogen and carbon atom transfer reactions to olefins and carbon–hydrogen bonds. Chapter 1 introduces directed evolution, a strategy for the laboratory optimization of proteins, in the context of improving metalloproteins for their native catalysis or for new reactions. Chapter 2 details the development of an enzyme-catalyzed transformation of olefins to aziridines, a valuable motif which is both present in bioactive molecules and used as a versatile building block for synthesis. This study establishes that when provided the appropriate reagents (e.g. styrenes and tosyl azide), heme proteins can adopt a nitrene transfer catalytic cycle to form aziridine products and that the turnover and selectivity of the catalyst can be optimized through mutation of its amino acid sequence. The activity of heme protein catalysts is extended to the functionalization of  $sp^3$  hybridized C–H bonds for carbon–nitrogen and carbon–carbon bond formation through nitrene and carbene insertion respectively (Chapters 3 and 4). With the exception of C–H oxygenation chemistry, iron complexes are under-utilized for  $sp^3$  C–H functionalization reactions, despite iron being readily available and non-toxic. Combining previously engineered heme proteins with suitable substrates led to initial reaction discovery. Directed evolution of these enzymes significantly improved their C–H functionalization activity (by 140-fold in Chapter 4). Characterization of evolved enzymes, including the attainment of an X-ray crystal structure (Chapter 3) and substrate scope studies (Chapters 3 and 4), were pursued. In sum, the thesis work addresses both the biological question of expanding the catalytic capabilities of existing enzymes through mutation and expands the chemistry of iron-porphyrin catalysts.



## PUBLISHED CONTENT AND CONTRIBUTIONS

(<sup>†</sup> indicates equal contributions by authors)

- (1) Farwell, C. C. <sup>†</sup>; **Zhang, R. K.** <sup>†</sup>; McIntosh, J. A.; Hyster, T. K.; Arnold F. H. “Enantioselective enzyme-catalyzed aziridination enabled by active-site evolution of a cytochrome P450,” *ACS Central Science* **2015**, *1*, 89–93. **DOI:** 10.1021/acscentsci.5b00056.

R. K. Z. participated in the design and execution of research, including protein engineering and substrate scope experiments. R. K. Z. participated in writing of the manuscript.

- (2) Prier, C. K. <sup>†</sup>; **Zhang, R. K.** <sup>†</sup>; Buller, A. R.; Brinkmann-Chen, S.; Arnold, F. H. “Enantioselective, intermolecular benzylic C–H amination catalysed by an engineered iron-haem enzyme,” *Nature Chemistry* **2017**, *9*, 629–634. **DOI:** 10.1038/nchem.2783.

R. K. Z. participated in the design and execution of research, including protein engineering, enzymatic reactions, and structural and mechanistic experiments. R. K. Z. participated in writing of the manuscript.

- (3) **Zhang, R. K.**; Romney, D. K.; Kan, S. B. J.; Arnold, F. H. “Chapter 5. Directed evolution of artificial metalloenzymes: bridging synthetic chemistry and biology” In: Dieguez, M.; Bäckvall, J.-E.; Pamies, O., editors. *Artificial Metalloenzymes and MetalloDNazymes in Catalysis. From Design to Applications*. Wiley-VCH **2018**, 137–170. **DOI:** 10.1002/9783527804085.ch5.

R. K. Z. prepared the review article.

- (4) **Zhang, R. K.**; Chen, K.; Huang, X.; Wohlschlager, L.; Renata, H.; Arnold, F. H. “Enzymatic assembly of carbon–carbon bonds via iron-catalysed  $sp^3$  C–H functionalisation,” *Nature* **2019**, *565*, 67–72. **DOI:** 10.1038/s41586-018-0808-5.

R. K. Z. participated in the design and execution of the research. R. K. Z. prepared the manuscript.

- (5) **Zhang, R. K.**; Huang, X.; Arnold, F. H. “Selective C–H bond functionalization with engineered heme proteins: New tools to generate complexity,” *Current Opinion in Chemical Biology* **2019**, *49*, 67–75. DOI: 10.1016/j.cbpa.2018.10.004.

R. K. Z. prepared the review article.

## PUBLISHED CONTENT NOT INCLUDED IN THESIS AND CONTRIBUTIONS

- (6) Herwig, L. <sup>†</sup>; Rice, A. J. <sup>†</sup>; Bedbrook, C. N.; **Zhang, R. K.**; Lignell, A.; Cahn, J. K. B.; Renata, H.; Dodani, S. C.; Cho, I.; Cai, L.; Gradinaru, V.; Arnold, F. H. “Directed evolution of a bright near-infrared fluorescent rhodopsin using a synthetic chromophore,” *Cell Chemical Biology* **2017**, *24*, 89–93. DOI: 10.1016/j.chembiol.2017.02.008.

R. K. Z. synthesized and characterized retinal analog used for biological experiments.

R. K. Z. participated in the preparation of the manuscript.

- (7) Chen, K.; Huang, X.; Kan, S. B. J.; **Zhang, R. K.**; Arnold, F. H. “Enzymatic construction of highly strained carbocycles,” *Science* **2018**, *360*, 71–75. DOI: 10.1126/science.aar4239.

R. K. Z. participated in enzyme experiments. R. K. Z. participated in reviewing and editing the manuscript.

- (8) Cho, I.; Prier, C. K.; Jia, Z.-J.; **Zhang, R. K.**; Görbe, T.; Arnold, F. H. Enantioselective aminohydroxylation of styrenyl olefins catalyzed by an engineered hemoprotein. *Angewandte Chemie International Edition* **2019**, just accepted. DOI: 10.1002/anie.201812968.

R. K. Z. participated in compound synthesis and data analysis. R. K. Z. participated in reviewing and editing the manuscript.

## TABLE OF CONTENTS

Acknowledgements.....	iv
Abstract .....	v
Published Content and Contributions.....	vi
Table of Contents.....	viii
List of Tables.....	x
List of Figures.....	xii
Abbreviations.....	xv
<b>Chapter I. Directed Evolution of Metalloproteins: An Introduction.....</b>	<b>1</b>
1.1 Opening Remarks.....	1
1.2 Strategies for Directed Evolution.....	2
1.3 Directed Evolution as an Uphill Walk in the Protein Fitness Landscape .....	4
1.4 Improving Enzymes for Native Reactions.....	5
1.5 Directed Evolution of Heme Proteins for Carbene and Nitrene Transfer Reactions.....	8
1.6 Optimization of Proteins with Artificial Metallocofactors or Metal-Binding Sites.....	12
1.7 Outlook.....	15
1.8 References and Notes.....	16
<b>Chapter II. Enantioselective Enzyme-Catalyzed Aziridination Enabled     by Evolution of a Cytochrome P450.....</b>	<b>22</b>
2.1 Abstract.....	22
2.2 Introduction.....	22
2.3 Results and Discussion.....	25
2.3.1 Identification of Heme Proteins with Promiscuous Aziridination Activity.....	25
2.3.2 Active-Site Engineering of an Enzyme Aziridination Catalyst.....	26
2.3.3 Substrate Scope.....	31
2.3.4 Tuning the Enantioselectivity of Enzyme-Catalyzed Olefin Aziridination .....	33
2.4 Conclusion.....	34

2.5 Experimental Methods.....	34
2.6 References and Notes.....	40
<b>Appendix A. Supplementary Information for Chapter 2.....</b>	<b>45</b>
<b>Chapter III. Intermolecular Benzylic C–H Amination Catalyzed by</b>	
Engineered Cytochrome P450 Enzymes.....	73
3.1 Abstract.....	73
3.2 Introduction.....	73
3.3 Results and Discussion.....	76
3.3.1 Reaction Discovery and Directed Evolution.....	76
3.3.2 <i>In Vitro</i> Studies.....	79
3.3.3 Structural Studies.....	82
3.3.4 Substrate Scope.....	85
3.3.5 Extended Substrate Scope Studies.....	88
3.4 Conclusion.....	90
3.5 Experimental Methods.....	90
3.6 References and Notes.....	99
<b>Appendix B. Supplementary Information for Chapter 3.....</b>	<b>105</b>
<b>Chapter IV. Carbon–Carbon Bond Formation by Heme Protein</b>	
Catalyzed C–H Functionalization.....	127
4.1 Abstract.....	127
4.2 Introduction.....	128
4.3 Results and Discussion.....	130
4.3.1 Reaction Discovery and Directed Evolution.....	130
4.3.2 Kinetic Isotope Effect Study.....	133
4.3.3 Benzylic C–H Insertion Substrate Scope.....	135
4.3.4 Enzymatic Functionalization of Allylic, Propargylic, and $\alpha$ -Amino C–H bonds .....	137
4.3.5 Diazo Reagent Scope.....	140
4.4 Conclusion.....	141
4.5 Experimental Methods.....	141
4.6 References and Notes.....	149
<b>Appendix C. Supplementary Information for Chapter 4.....</b>	<b>153</b>

## LIST OF TABLES

<i>Number</i>	<i>Page</i>
1-1 Improvement of P450 <sub>BM3</sub> for a non-natural cyclopropanation Reaction.....	10
2-1 Total turnovers to product for aziridination catalyzed by purified holoenzymes P411 <sub>BM3</sub> -CIS T438S (P) and P-I263F.....	27
2-2 Improvement in yield and % <i>ee</i> for aziridine product <b>19</b> with active-site evolution of P411 <sub>BM3</sub> CIS T438S (P) .....	28
2-3 Initial rates of aziridination and azide reduction for engineered enzymes.....	31
2-4 Substrate scope of aziridination with P-I263F A328V L437V.....	32
A-1 Mutations present in P450 <sub>BM3</sub> variants used in Chapter 2.....	45
A-2 Heme and other hemeproteins tested for aziridination activity.....	46
A-3 Panel of P450 <sub>BM3</sub> and P411 purified enzymes tested for aziridination activity .....	47
A-4 Control experiments using purified P-I263F proteins and NADPH as reductant .....	48
3-1 C–H amination of performed with purified P411 variants .....	81
3-2 Initial rates of C–H amination and azide reduction .....	81
3-3 Optimization of the P11 <sub>CHA</sub> -catalyzed amination of 4-ethylanisole .....	86
3-4 Initial tests for C–H amidation using reagent <b>48</b> with select cytochrome P411 variants .....	89
3-5 Initial tests for C–H amidation using reagent <b>49</b> with select cytochrome P411 variants .....	89
3-6 X-ray crystallography analysis of cytochrome P411 P-4 A82L A78V F263L heme domain (PDB: 5UCW). ....	98
B-1 C–H amination of 4-ethylanisole with select variants of cytochrome P450 <sub>BM3</sub> .....	105
B-2 C–H amination of 4-ethylanisole performed with hemein or select heme proteins. ....	107

B-3	C–H amination with P411 variants previously engineered for the imidation of allylic sulfides .....	108
B-4	Mutations present in cytochrome P450 <sub>BM3</sub> variants used in Chapter 3.....	108
B-5	Summary of directed evolution for intermolecular C–H amination.....	109
B-6	Data to accompany Figure 3-5 .....	110
B-6	Demonstration that amination of isochroman with cytochrome P411 <sub>CHA</sub> is enantioselective.....	111
C-1	Initial results for C–H alkylation with <b>51a</b> and ethyl diazoacetate using heme proteins and control reactions.....	153
C-2	Initial results for C–H alkylation with 4-ethyl anisole and ethyl diazoacetate using heme proteins and control reactions.....	154
C-3	A P411ΔFAD variant is functional in whole E. coli cells, lysate, and <i>in vitro</i> .....	154
C-4	Enzymatic reactions under non-standard conditions.....	155
C-5	Preliminary experiments with heme proteins.....	166
C-6	Heme proteins tested for C–H alkylation activity.....	167
C-7	Engineered cytochrome P411 variants show promiscuous C–H alkylation activity.....	168
C-8	Summary of directed evolution for C–H alkylation.....	170
C-9	Mutations and truncations relative to P450 <sub>BM3</sub> wild-type for the P411 variants described in Table C-8.....	172
C-10	Other P411 variants described in this study.....	173
C-11	Enzymatic C–H alkylation data presented in Figure 4-4.....	174
C-12	Enzymatic C–H alkylation data presented in Figure C-1.....	175

## LIST OF FIGURES

<i>Number</i>	<i>Page</i>
1-1 Overview of directed evolution.....	3
1-2 Hypothetical protein fitness landscape.....	5
1-3 Directed evolution of a carbonic anhydrase.....	6
1-4 Regiodivergent hydroxylation of testosterone by engineered P450 <sub>BM3</sub> variants.....	7
1-5 Regiodivergent C–H amination catalyzed by engineered P450 <sub>BM3</sub> variants .....	11
1-6 Carbene transfer catalyzed by metal-substituted myoglobin variant .....	12
1-7 Artificial Zn-binding protein of Tezcan and co-workers .....	13
1-8 Examples of hybrid catalysts based on the biotin-streptavidin technology.....	14
1-9 A platform for Rh-protein conjugates based on proline peptidase .....	15
2-1 Reactions of aziridines, select examples.....	23
2-2 Strategies to access aziridines.....	23
2-3 Natural biosynthetic routes to aziridines.....	24
2-4 Initial olefin aziridination reactivity catalyzed by purified holoenzyme P411 <sub>BM3</sub> -CIS T438S .....	26
2-5 Comparison of active site residues in P411 <sub>BM3</sub> -CIS and P-I263F.....	29
2-6 Proposed mechanism of a cytochrome P411-catalyzed aziridination reaction .....	30
2-7 Enzyme-controlled enantioselectivity of olefin aziridination .....	33
A-1 Demonstration of enzymatic synthesis and degradation of aziridine <b>16'</b> under reaction conditions.....	51
A-2 UV-vis absorbance spectra of P-I263F (full-length) and P-I263F heme domain only proteins after addition of NADPH.....	52
A-3 Initial rate data to accompany Table A-5.....	54

A-4	Assignment of absolute stereochemistry of enzymatically produced aziridine <b>21</b> by chiral HPLC.....	59
3-1	Examples of chiral amine containing pharmaceutical agents.....	74
3-2	Intermolecular C–H amination overview.....	75
3-3	Proposed mechanism of intermolecular C–H amination catalyzed by a cytochrome P411 enzyme.....	76
3-4	Comparison of sulfide imidation and C–H amination substrate geometries.....	77
3-5	Evolution of a cytochrome P411 catalyst for enantioselective C–H amination.....	79
3-6	Effect of NADPH concentration on C–H amination activity of P411 <sub>CHA</sub> <i>in vitro</i> .....	80
3-7	Kinetic isotope effect studies with P411 <sub>CHA</sub> .....	82
3-8	Active site view of the P-4 A82L A78V F263L structure.....	83
3-9	Heme domain protein alignments of cytochrome P411 variant P-4 A82L A78V F263L .....	84
3-10	Root-mean-square deviation analysis P411 structures .....	85
3-11	Substrate scope of enzymatic intermolecular C–H amination .....	87
B-1	Initial rates of enzymatic C–H amination of 4-ethylanisole.....	113
B-2	Initial rates of azide reduction in the enzymatic C–H amination of 4-ethylanisole.....	114
B-3	Independent rate experiments for kinetic isotope effect determination.....	115
4-1	Enzymatic C–H functionalization systems.....	129
4-2	Select subset of heme proteins tested for promiscuous <i>sp</i> <sup>3</sup> C–H alkylation activity. ....	131
4-3	Truncation of a full-length P411 protein delivers active C–H alkylation enzymes.....	132
4-4	Directed evolution of a cytochrome P411 for enantioselective C–H alkylation.....	133
4-5	Kinetic isotope effect of C–H alkylation by P411-CHF.....	134
4-6	Substrate scope for benzylic C–H alkylation with P411-CHF.....	136



4-7	Application of P411 enzymes for $sp^3$ C–H alkylation.....	139
4-8	Enzymatic C–H alkylation with alternative diazo reagents.....	140
C-1	Evolutionary lineage from P-4 A82L to P411-CHF evaluated for C–H alkylation of 4-ethylanisole.....	156
C-2	Structural visualization of amino acids mutated during directed evolution.....	157
C-3	Structural visualization of amino acid differences between P450 <sub>BM3</sub> wild-type and P411-CHF.....	158
C-4	Product profiles for the reaction of 4-allylanisole and ethyl diazoacetate with P411-CHF and with P-I263F.....	159
C-5	GC traces of enzymatic reactions of <b>51h</b> with ethyl diazoacetate.....	160
C-6	GC trace of enzymatic reaction of <b>53e</b> with ethyl diazoacetate.....	161
C-7	GC trace of enzymatic reaction of <b>56a</b> with ethyl diazoacetate.....	162
C-8	HPLC trace of enzymatic reaction of <b>56b</b> with ethyl diazoacetate.....	163
C-9	Additional diazo substrates tested.....	164
C-10	Alkane substrates used in Ch. 4.....	177
C-11	Diazo compounds used in Ch. 4.....	178

## ABBREVIATIONS

$[\alpha]_D$	angle of optical rotation of plane-polarized light
Å	angstrom(s)
Ac	acetyl
Ar	aryl group
BM3	cytochrome P450 from <i>Bacillus megaterium</i> (CYP102A1)
<i>c</i>	concentration of sample for measurement of optical rotation
CO	carbon monoxide
CYP	cytochrome P450 monooxygenase identifier
DMSO	dimethyl sulfoxide
d.r.	diastereomeric ratio
ee	enantiomeric excess
e.r.	enantiomeric ratio
Et	ethyl
EtOH	ethanol
GC	gas chromatography
GC-MS	gas chromatography with mass spectrometry
h	hour(s)
HB <sub>amp</sub>	Hyperbroth medium with 0.1 mg/mL ampicillin
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
<i>k</i>	rate constant
KPi	potassium phosphate
LB <sub>amp</sub>	Luria-Bertani medium with 0.1 mg/mL ampicillin
LC-MS	liquid chromatography with mass spectrometry
M9-N	M9 minimal media with no nitrogen source added
Me	methyl
NADPH	reduced nicotine adenine dinucleotide phosphate
P450	cysteine-ligated cytochrome P450 monooxygenase
P411	serine-ligated cytochrome P450 enzyme
PDB	Protein Data Bank
r.r.	regiomeric ratio
RT	room temperature
t <sub>R</sub>	retention time
Ts	para-toluenesulfonyl (tosyl)

## DIRECTED EVOLUTION OF METALLOPROTEINS: AN INTRODUCTION

Material for this chapter appears in **Zhang, R. K.**; Romney, D. K.; Kan, S. B. J.; Arnold, F. H. “Chapter 5. Directed evolution of artificial metalloenzymes: bridging synthetic chemistry and biology” In: Dieguez, M.; Bäckvall, J.-E.; Pamies, O., editors. *Artificial Metalloenzymes and MetalloDNazymes in Catalysis. From Design to Applications*. Wiley-VCH **2018**, 137–170. **DOI:** 10.1002/9783527804085.ch5. Reprinted with permission from John Wiley and Sons.

Related material:

**Zhang, R. K.**; Huang, X.; Arnold, F. H. “Selective C–H bond functionalization with engineered heme proteins: New tools to generate complexity,” *Current Opinion in Chemical Biology* **2019**, 49, 67–75. **DOI:** 10.1016/j.cbpa.2018.10.004.

### 1.1 Opening Remarks

Nature is an expert at creating protein frameworks which can be repurposed to perform a multitude of chemical transformations. For example, the P450 superfamily consists of structurally similar heme-containing proteins that catalyze C–H oxygenation, alkene epoxidation, oxidative cyclization,<sup>1</sup> aryl-aryl coupling,<sup>2</sup> and nitration,<sup>3</sup> among others.<sup>4</sup> While an enzyme may be optimized to catalyze a specific reaction on a native substrate, it often exhibits activity on alternative substrates or may even be able to promote a different transformation. Evolution takes advantage of these ‘promiscuous’ activities in response to a changing environment to create and optimize enzymes for new functions.<sup>5–7</sup>

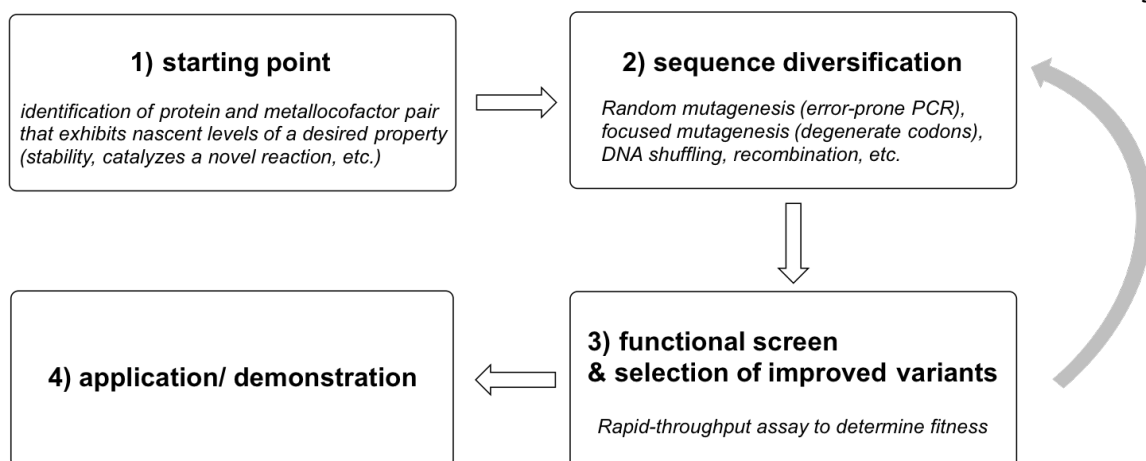
Directed evolution, the laboratory mimic of adaptive evolution, harnesses the malleability of proteins to invent useful catalysts suitable to the experimenter’s needs. In this process, beneficial mutations identified by screening libraries are accumulated, one or a few at a time, to enhance desired function(s) or performance. This technique has been used to confer useful properties such as stability under harsh conditions and improved activity on non-native substrates (*vide infra*).

More recently, scientists have discovered that synthetically useful reactions not found in nature can be catalyzed by metalloproteins, designed or natural, and that these non-natural activities can also be improved by directed evolution. New activities can arise when natural metalloproteins are exposed to abiological reagents or when an artificial metallocofactor is introduced into a protein scaffold (*vide infra*). Indeed, this development has led to a paradigm shift in which proteins are no longer relegated to their annotated functions but can now participate as elements for designing synthetic methodology. Guided by a chemical perspective, target reactions can be investigated using suitable natural and designed proteins. Once even a low level of the activity has been found, the enzyme can be optimized and diversified by directed evolution.

This chapter will introduce directed evolution in the context of metalloproteins. Approaches that have been used to improve different enzyme systems are described in order to illustrate opportunities and requirements for using directed evolution. Finally, the thesis research is summarized with an emphasis on its contribution to conceptual advances in the field.

## **1.2 Strategies for Directed Evolution**

The goal of a directed evolution experiment is to imbue a protein with a new property, such as high selectivity, activity with a non-native substrate, or stability in a different temperature or pH range. The process involves four key steps: 1) identification of a starting protein (that may have only low levels of the desired property), 2) diversification of the starting protein through mutagenesis of the corresponding DNA sequence, 3) functional screening of the variants to identify improved variants, and 4) repeating steps 2 and 3 until sufficient levels of the desired property are achieved (Figure 1-1). With advances in molecular biology and analytical tools, there are many methods available for executing each step, the choice of which will affect the evolutionary trajectory and end point.



**Figure 1-1.** Overview of directed evolution.

Identification of a starting protein is the most crucial step of this process. One requirement is that the starting protein exhibits a measurable level of the desired function. In the case for catalysis, this would mean that the starting protein needs to have a measurable level of activity for the desired transformation under target conditions. Knowledge of chemical mechanisms, both of the target reaction and natural enzyme mechanisms, can aid in the choice of suitable proteins to test or design. Other properties of a protein, such as thermostability<sup>8</sup> and the plasticity of the fold in natural enzyme diversity,<sup>4</sup> may impact the evolvability of a chosen starting point.

Screening is one of the most experimentally challenging steps in directed evolution of synthetically useful enzymes, and screening capacity will often determine how the sequence diversification is done. Because the goal of the experiment is to find beneficial mutations, the variant library should have a frequency of beneficial mutations that matches the screening capability. Of course, the beneficial mutation frequency is not known at the beginning, but one can make educated guesses. In random mutagenesis, mutations are made throughout the protein. This approach is advantageous because it recognizes that the effects of mutations on protein structure, dynamics, and catalysis are not predictable *a priori*. For instance, mutations distant from an enzyme's active sites can affect catalysis.<sup>9</sup>

<sup>10</sup> However, even for a small protein, the number of possible ways to make even a single

mutation is so large that a rapid-throughput screen, such as a colorimetric assay, is usually required to sample enough sequences to find a beneficial mutation. (Whole protein random mutagenesis at 1–3 mutations per gene usually requires screening hundreds to thousands of mutant proteins per generation.<sup>11</sup>) Unfortunately, many desired properties are not easily assayed in high throughput.

An alternative is to introduce mutations in a focused manner, guided by structural, mechanistic, or computational information. The assumption is that beneficial mutations will appear at the chosen positions (e.g., residues in the active site) with a higher frequency than at positions chosen at random.<sup>12</sup> With the use of degenerate codons, focused libraries of variants can be created in which one or more specific residues are mutated to all or selected subsets of the 20 proteinogenic amino acids.<sup>13</sup> This approach can significantly reduce library size and focus on amino acids believed to be key. However, success rests on making the right choice of residues to target. The fact remains that the effect of a mutation is often unpredictable. Thus, sites hypothesized to be important may in fact exert little influence on the desired function or property, or may not even tolerate mutation.

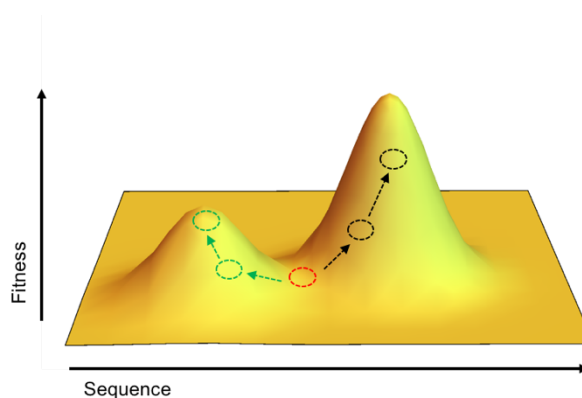
Other methods for sequence diversification include recombination, wherein genes or gene fragments of related proteins are shuffled and reassembled, thus enabling beneficial combinations of amino acids to appear in a single protein.<sup>14</sup> Furthermore, beneficial mutations that are found separately can be recombined in a combinatorial fashion, or hybrids of evolutionarily related proteins can be made. Frequently, a combination of diversification methods is used, especially if multiple rounds of directed evolution are necessary.<sup>15–17</sup>

### **1.3 Directed Evolution as an Uphill Walk in the Protein Fitness Landscape**

The concept of directed evolution as a walk through sequence space that passes through functional proteins was born out of a paper by John Maynard Smith in 1970.<sup>18</sup> Smith asserted that in order for evolution by natural selection to be successful, there must be a continuous trail of functional proteins that can be traversed in single mutational steps

without passing through nonfunctional proteins. He illustrated the idea with a word game, using conversion of the word ‘WORD’ to ‘GENE’ one letter at a time, with the requirement that all intermediate words be meaningful. A nonsense word, analogous to a nonfunctional protein, would be eliminated by selection before a second change could occur. His solution to the game, WORD → WORE → GORE → GONE → GENE, interconverts the two words through a series of single letter changes, each of which produces a meaningful word.

Directed evolution takes the game one step further and searches for improvements. With Smith’s analogy, we can assign each word (and the protein it represents) a fitness, defined as the number of letters it matches with ‘GENE’, and think of this transformation from ‘WORD’ to ‘GENE’ as an adaptive uphill walk to a fitness peak. Laboratory protein evolution can then be envisioned as a walk on a high-dimensional fitness landscape in sequence space, where fitness is defined by the experimenter. Iterative rounds of mutagenesis and screening lead the experimenter to higher fitness levels, provided that the landscape is smooth (Figure 1-2).<sup>15</sup> We do not know the structure of a protein fitness landscape; indeed, we expect the structure of the landscape to depend on the specific protein and problem. Past directed evolution successes and the success of natural evolution, however, suggest that enzyme landscapes are smooth for some properties, including catalysis and stability (*vide infra*), in at least some sequence dimensions.



**Figure 1-2.** Hypothetical protein fitness landscape. Directed evolution takes a protein along its fitness landscape, where fitness is a metric defined by the experimenter. Sequence diversification samples the nearby sequences, and screening identifies fitness improvements. Two possible evolutionary trajectories (green and black) from a single

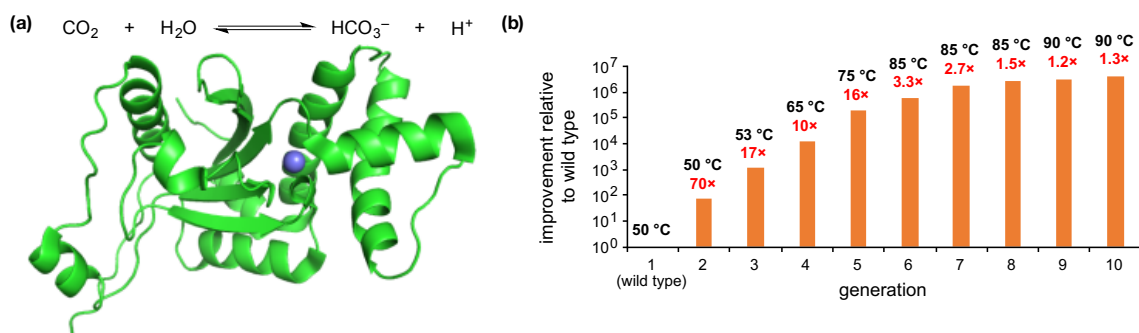
starting point (red) illustrate that there may exist multiple local maxima or solutions to any given optimization, and the results may depend on the path taken. Sequence space is of very high dimensionality; a simplified fitness landscape is presented here.

## 1.4 Improving Enzymes for Native Reactions

Natural metalloenzymes catalyze a myriad of chemical transformations. However, the utility of these biocatalysts for chemical synthesis or industrial applications is limited. This could be due to restricted substrate scope or instability under process conditions. Directed evolution can address these issues by altering the activity profile of an enzyme or changing the range of optimal operating conditions. Here, recent examples with carbonic anhydrase and cytochrome P450 are highlighted to illustrate the success of this approach for enhancing biocatalyst performance on a native reaction.

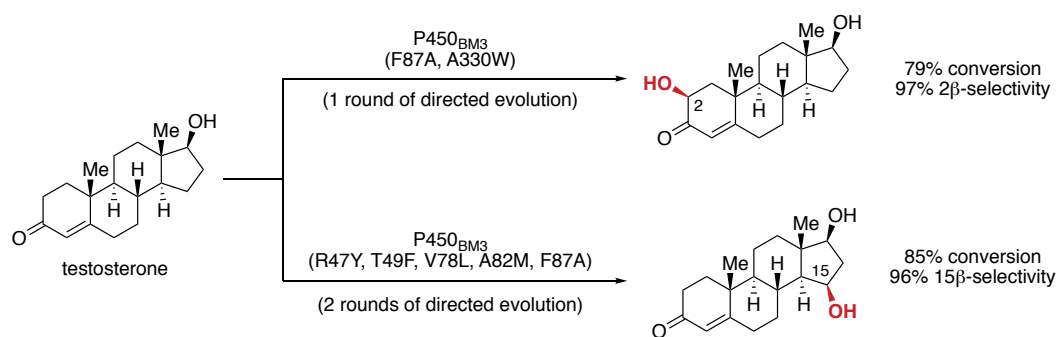
Carbonic anhydrase is a Zn-dependent metalloenzyme that catalyzes the reversible hydration of carbon dioxide into bicarbonate and a proton (Figure 1-3a). It is potentially a useful catalyst for carbon capture and sequestration (CCS), a process to reduce anthropogenic CO<sub>2</sub> released from fossil fuel emissions.<sup>19</sup> However, naturally occurring carbonic anhydrases are limited by their inability to tolerate the harsh alkaline solvent and high process temperatures required for CCS. This challenge was addressed by Lalonde and coworkers at Codexis, who used directed evolution to generate a highly stable carbonic anhydrase capable of withstanding CO<sub>2</sub> capture conditions.<sup>20</sup> Nine rounds of evolution converted the parent enzyme (wild type carbonic anhydrase from *Desulfovibrio vulgaris*), which had negligible activity after exposure to aqueous amine solvent at 50 °C, to a variant that retained activity after heating to 107 °C in alkaline solvent (Figure 1-3b). Mutations at 36 positions (representing 15% of the enzyme) were accumulated during the evolution. Exactly how the mutations enhance protein stability is not known, and this optimization solution could not be predicted *a priori*. The utility of an ultrastable variant was demonstrated in a pilot scale CO<sub>2</sub> capture system. Comparison of pilot scale CO<sub>2</sub> capture demonstrations performed in the presence or absence of the ultrastable enzyme showed that the rate of CO<sub>2</sub> capture was enhanced 25-fold using the evolved variant.





**Figure 1-3.** Directed evolution of a carbonic anhydrase stable to CO<sub>2</sub> capture process conditions.<sup>20</sup> **(a)** Carbonic anhydrase catalyzes the reversible hydration of CO<sub>2</sub> to bicarbonate and a proton. The enzyme utilizes a catalytic zinc atom, displayed as a sphere. Only a single subunit of the tetrameric protein is shown (PDB 2A5V). **(b)** Evolution of an ultrastable carbonic anhydrase. Half-lives of the variants at the indicated temperature (black) were determined by measuring CO<sub>2</sub> absorption in a reactor. The fold improvement over the previous round is shown above each bar (red).

The cytochrome P450 superfamily has exploited a versatile protein framework to create enzymes with many different functions. A good indication of a protein family's evolvability in the laboratory is its functional diversity in nature,<sup>21</sup> and indeed P450s are among the most studied metalloenzymes for directed evolution. In one example, two variants derived from cytochrome P450<sub>BM3</sub> from *Bacillus megaterium* were engineered for regiodivergent hydroxylation of a non-native steroid substrate, testosterone.<sup>22</sup> Although testosterone is not accepted as a substrate by wild type P450<sub>BM3</sub>, a single mutation activated the enzyme toward this substrate. Using degenerate codons, sets of amino acid residues were mutated simultaneously and screened for improved selectivity for oxidation at the 2β- or 15β-position (Figure 1-4). Molecular dynamics simulations and substrate docking were performed on a non-selective variant, a 2β-selective variant, and a 15β-selective variant to investigate the basis of the observed regioselectivity. The results of these studies suggest that in the non-selective variant, testosterone can adopt two energetically equivalent orientations, one of which would lead to the 2β-hydroxylated product and the other to the 15β-hydroxylated product. In contrast, for the 2β-selective and the 15β-selective enzymes, each has a distinct binding pocket that allows only one active orientation of testosterone.



**Figure 1-4.** Regiodivergent hydroxylation of testosterone by engineered P450<sub>BM3</sub> variants.<sup>22</sup>

Additional applications of directed evolution to expand the scope of P450 oxidation reactions have led to the identification of numerous variants with impressive activities.<sup>23</sup> These activities include hydroxylation of simple molecules, such as short-chain alkanes,<sup>24–27</sup> as well as complex molecules, such as anti-malarial therapeutic artemisinin<sup>28</sup> and natural product parthenolide,<sup>29</sup> with defined regio- and stereoselectivities. Evolution for chemoselectivity has also been investigated; in one report, an engineered variant displayed up to 90% chemoselectivity in favor of allylic hydroxylation versus epoxidation.<sup>30</sup> As collections of diverse cytochrome P450 variants expand, more applications for these malleable protein scaffolds will undoubtedly arise.

## 1.5 Directed evolution of heme proteins for carbene and nitrene transfer reactions

Although metalloproteins catalyze an impressive set of reactions, there are many synthetically useful metal-catalyzed transformations that are *not* found in nature—the question becomes: how do we identify proteins capable of acquiring some of these non-natural activities? One approach to creating new enzymes relies on using the catalytic promiscuity of natural enzymes to find suitable starting points for directed evolution. Metalloproteins containing versatile cofactors are a good place to start. In mechanistically guided efforts, researchers challenge these proteins with synthetic reagents known to be capable of reacting with the cofactor. This approach has been especially successful with heme proteins, which have served as starting points for a number of synthetically useful

carbene and nitrene transfer reactions in recent years.<sup>31</sup> A second strategy is to replace the cofactor with a structurally similar derivative. This has been demonstrated through the creation of catalytically active heme proteins in which the heme prosthetic group is replaced with an Ir(Me)-porphyrin (*vide infra*), among others.<sup>32,33</sup> After identifying a starting point, the next question is whether the fitness landscape for these abiological reactions is locally smooth and contains higher peaks, such that the new activity can be optimized through directed evolution. The answer appears to be yes, natural metalloproteins and their derivatives can acquire new, non-natural activities, and their performance can be enhanced by directed evolution. Seminal examples are discussed in the following paragraphs.

Iron-porphyrin complexes have been used as models to study the reaction mechanism of cytochrome P450s.<sup>34</sup> Since then, these complexes have been shown to catalyze a variety of reactions, some of which are not found in nature. For example, the reaction of an iron-porphyrin complex with a diazo compound and an alkene produces a cyclopropane product, with the reaction thought to proceed through a reactive iron-carbene intermediate.<sup>35</sup> Other examples of carbene transfer reactions mediated by iron porphyrins include carbene insertion into N–H bonds<sup>36</sup> and formation of phosphonium ylides.<sup>37</sup>

Given this body of work and the similarity between the iron-carbene and iron-oxo reactive intermediate in the natural P450 hydroxylation mechanism, Coelho *et al.* hypothesized that a cytochrome P450 could perform abiological carbene transfer reactions.<sup>38</sup> Indeed, the authors found that wild type cytochrome P450<sub>BM3</sub> and variants of this enzyme, as well as other heme proteins, had this promiscuous activity. Using styrene and ethyl diazoacetate as model substrates, they demonstrated that P450<sub>BM3</sub> could be optimized by mutagenesis and screening to become an efficient and highly selective cyclopropanation catalyst (Table 1-1). Since this report, other carbene transfer enzymes derived from the cytochrome P450<sub>BM3</sub> scaffold have been developed; these enzymes catalyze reactions including cyclopropanation with other olefins<sup>39</sup> and alkyne substrates,<sup>40</sup> N–H insertion,<sup>41</sup> and C–H insertion (Chapter 5). Notably, an enzyme scaffold is not

required for carbene transfer. Catalysts derived from myoglobin, cytochrome *c*, and globin homologs, heme proteins with no known natural catalytic function, can also be evolved to become efficient and selective carbene transfer enzymes for reactions including olefin cyclopropanation,<sup>42</sup> Si–H insertion,<sup>43</sup> and B–H insertion.<sup>44</sup>

**Table 1-1.** Improvement of P450<sub>BM3</sub> for a non-natural cyclopropanation reaction.<sup>38</sup> TTN, total turnover number, is defined as the amount of indicated product divided by protein concentration.

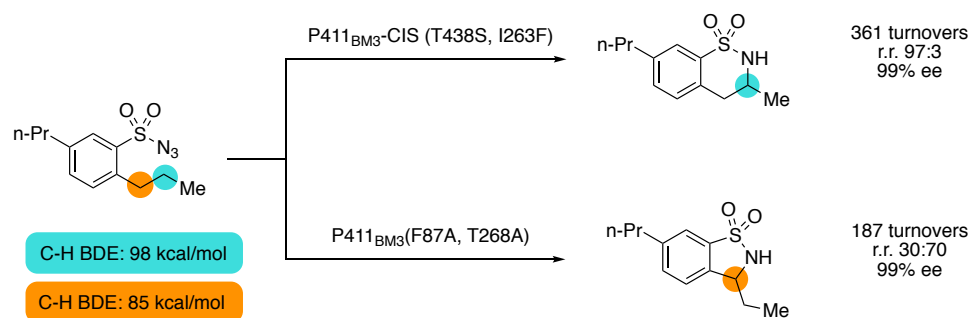
styrene + ethyl diazoacetate  $\xrightarrow[\text{phosphate buffer (0.1 M, pH 8.0), 5\% MeOH}]{\text{0.2 mol\% catalyst, 10 mM Na}_2\text{S}_2\text{O}_4}$  cis + trans

Entry	Catalyst	TTN	cis:trans	%ee <sub>cis</sub>	%ee <sub>trans</sub>
1	Hemin	73	6:94	1	0
2	P450 <sub>BM3</sub>	5	37:63	27	2
3	9-10A-TS(F87V)	7	35:65	41	8
4	P450 <sub>BM3</sub> -CIS	199	71:29	94	91
5	P450 <sub>BM3</sub> -CIS(T438S)	293	92:8	97	66

Directed Evolution

When given the opportunity to interact with certain nitrogen-containing substrates, heme proteins can putatively form iron-nitrene species and perform nitrogen-atom transfer reactions, such as C–H amination. Though first demonstrated in 1985,<sup>45</sup> it was only with the advent of modern directed evolution techniques that variants of P450<sub>BM3</sub> adopted the C–H amination function with synthetically useful levels of activity.<sup>46,47</sup> Impressively, the P450<sub>BM3</sub> protein scaffold could be engineered to provide strong control over substrate orientation during a non-natural C–H amination reaction, such that it could override the kinetic bias of the reaction.<sup>48</sup> This was demonstrated by developing two P450<sub>BM3</sub> variants with divergent regioselectivity for benzylic or homobenzylic C–H amination (Figure 1-5). Further work with variants of P450<sub>BM3</sub> and myoglobin have led to other catalysts for intramolecular C–H amination using arylsulfonyl azides as well as intermolecular aziridination (Chapter 2), sulfimidation,<sup>49</sup> and C–H amination (Chapter 3) with tosyl azide. Recent advances have also extended the scope of the nitrene precursor to include

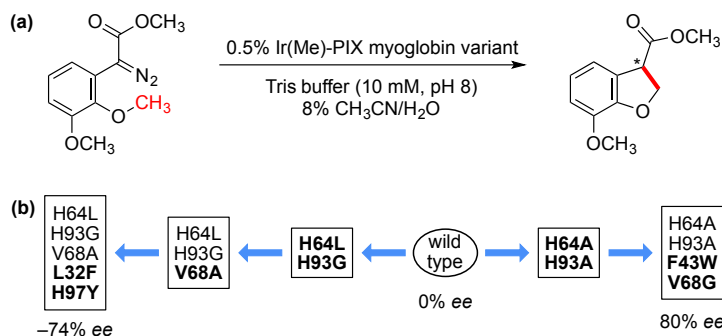
PivONH<sub>3</sub>OTf, which relies on the cleavage of an N–O bond for nitrene formation in contrast to the azide reagents which release N<sub>2</sub>. A method for aminohydroxylation of styrenes catalyzed by engineered variants of cytochrome *c* has been developed using the PivONH<sub>3</sub>OTf reagent.<sup>50</sup>



**Figure 1-5.** Regiodivergent C–H amination catalyzed by engineered cytochrome P450<sub>BM3</sub> variants.<sup>48</sup>

In an effort to further expand the reaction space accessible to heme proteins, replacement of the heme cofactor with other metal-porphyrins has been investigated. Sperm whale myoglobin and Cyp119, a P450 from *Sulfolobus acidocaldarius*, have been used as scaffolds for non-natural catalysis by replacing the native Fe of the heme cofactor with alternative metals. This approach is advantageous because it introduces the reactivity of abiological metals while retaining a metalloprotein scaffold that has already been shown to be evolvable. In the first report, it was found that myoglobin containing Ir(Me)-mesoporphyrin IX (Ir(Me)-PIX) in place of heme was a superior catalyst for an intramolecular carbene C–H insertion reaction compared to myoglobin with other metallated porphyrins (Fe, Co, Cu, Mn, Rh, Ru(CO), or Ag). Although the wild type Ir(Me)-PIX myoglobin was not enantioselective, variants with divergent enantioselectivity could be created through directed evolution (Figure 1-6).<sup>51</sup> Notably, the free Ir(Me)-PIX complex catalyzes the model reaction and with a higher reaction rate than evolved Ir(Me)-PIX myoglobin enzymes. Changing the protein scaffold to apo-CYP119 and subsequent directed evolution of the Ir(Me)-PIX Cyp119 protein delivered an artificial metalloenzyme with 23-fold higher turnover frequency compared with the free cofactor.<sup>52</sup> Other reactions,

including cyclopropanation<sup>53</sup> and intramolecular C–H amination<sup>54</sup>, have also been achieved using Ir(Me)-PIX proteins.



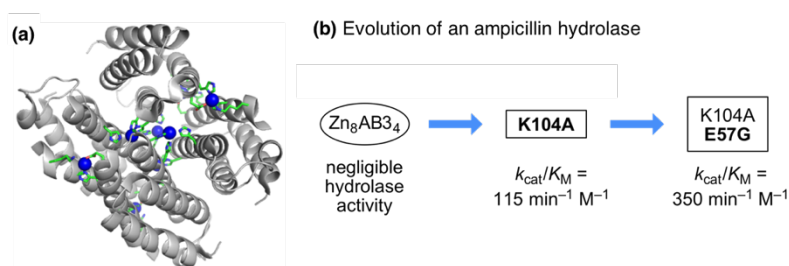
**Figure 1-6.** Carbene transfer catalyzed by metal-substituted myoglobin variant.<sup>51</sup> (a) Representative C–H insertion reaction. (b) Evolutionary trajectory resulting in variants which deliver the product shown in (a) with divergent enantioselectivity.

## 1.6 Optimization of Proteins with Artificial Metallocofactors or Metal-Binding Sites

Natural proteins cannot access the diversity of chemical transformations available to synthetic chemistry. This is in part because biological systems appear to use only a limited number of cofactors and elements from the periodic table. Scientists have sought to expand the catalytic repertoire of enzymes by creating proteins with completely artificial metallocofactors. One approach for the creation of such proteins is to introduce amino acids that can coordinate metal ions through their side-chains (e.g., histidine, cysteine, lysine). Several properly positioned coordinating residues will create a new metal-binding site that can compete for locally present metal ions. A second approach is to functionalize an organometallic complex with a pendant moiety that can bind to a protein scaffold. While both approaches have been used to create a number of novel metalloproteins, there are only a few examples of such proteins being optimized for catalytic activity through directed evolution.

In one example, directed evolution was applied to homo-tetrameric proteins with three-coordinate Zn(II) sites competent for ampicillin hydrolysis to improve this catalytic activity.<sup>55</sup> The homo-tetrameric proteins were derived from modified cytochrome *cb*<sub>562</sub>

building blocks which spontaneously self-assemble through four-coordinate structural Zn(II) coordination, hydrophobic interactions, and disulfide bonds. They were designed to have triads of histidine and glutamate residues that could sequester Zn(II) ions which would impart the hydrolysis activity. Although the initial construct, which contains eight zinc ions (Zn<sub>8</sub>:AB3<sub>4</sub>, Figure 1-7a), exhibited no hydrolase activity, a single mutation (K104A) led to a variant with detectable activity against ampicillin (Figure 1-7b). Furthermore, its hydrolase activity with ampicillin was sufficient to enable *Escherichia coli* to survive on ampicillin-containing media, thus allowing the authors to use cell survival as a metric for ampicillin-hydrolysis activity. Site-saturation mutagenesis at four positions around the active zinc ions, performed on the K104A variant, identified a second mutation that improved catalytic activity.

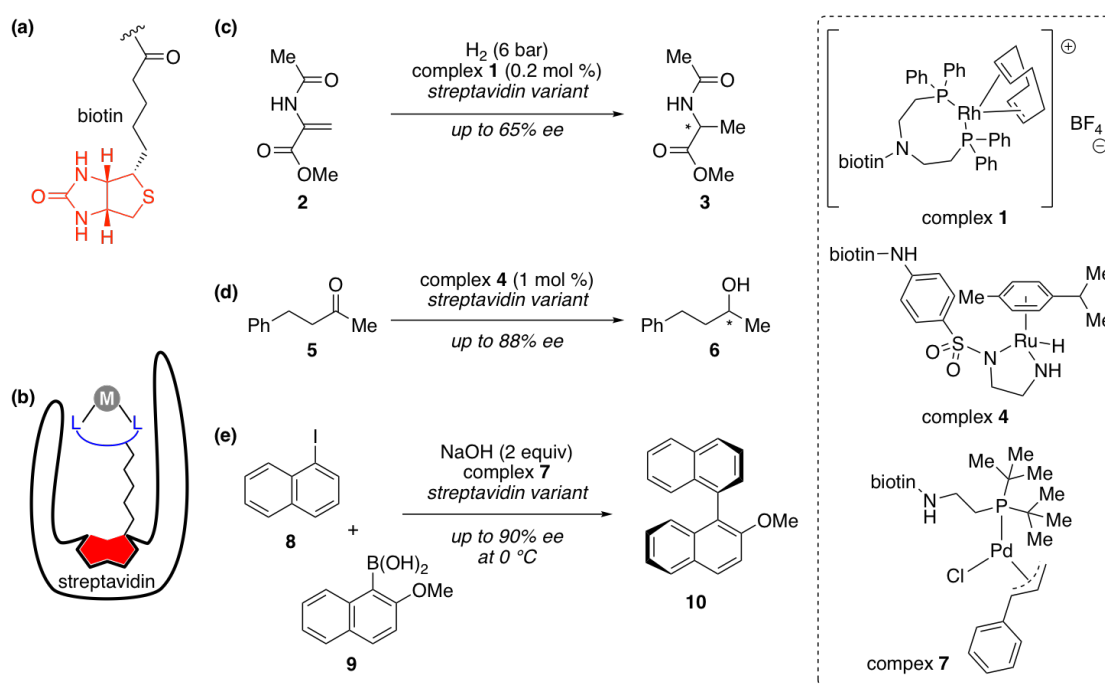


**Figure 1-7.** Artificial Zn-binding protein of Tezcan and co-workers.<sup>55</sup> **(a)** Tetrameric complex that served as parent for directed evolution (PDB 4U9D). **(b)** Progression of evolution for ampicillin-hydrolase activity.

In addition to metal ions, synthetic chemists have created a vast array of organometallic catalysts which could be conjugated to proteins. In principle, any of these catalysts can be conjugated to a protein to create a hybrid catalyst whose properties may be tuned through directed evolution. In practice, however, in order for these hybrid catalysts to achieve high activity and selectivity, the metal center must be sufficiently surrounded by the protein and held in a well-defined orientation so that the protein may exert significant influence on the reaction.

One conjugation strategy explored for construction of artificial metalloenzymes is to append a biotin linker (Figure 1-8a) to an organometallic complex, and then incorporate it into a biotin-binding protein, typically streptavidin (Figure 1-8b). Complexes which have

been incorporated into streptavidin and the target reactions of these hybrid catalysts include Rh-complex **1** for hydrogenation of acrylate **2** to alanine derivative **3**,<sup>56</sup> Ru-complex **4** for reduction of ketones to alcohols,<sup>57</sup> and Pd-complex **7** for Suzuki cross-coupling (Figure 1-8).<sup>58</sup> In these examples, a catalytically active organometallic catalyst was introduced into the streptavidin scaffold with the hypothesis that the protein would impart selectivity to the reaction. Mutagenesis to the streptavidin protein and screening of target reactions was performed to improve enantioselectivity. In contrast to the examples in Figure 1-8, which required purification of streptavidin, Ward and co-workers recently demonstrated that a streptavidin-based hybrid catalyst with a biotinylated Hoveyda-Grubbs second-generation complex can catalyze alkene metathesis in the presence of whole cells (not shown).<sup>59</sup> This advance was enabled by targeting streptavidin to the periplasm of *E. coli* cells. Since the periplasm is an oxidizing environment the majority of glutathione, an inhibitor of the hybrid catalyst, is in its oxidized form disulfide form which does not impact metathesis activity.

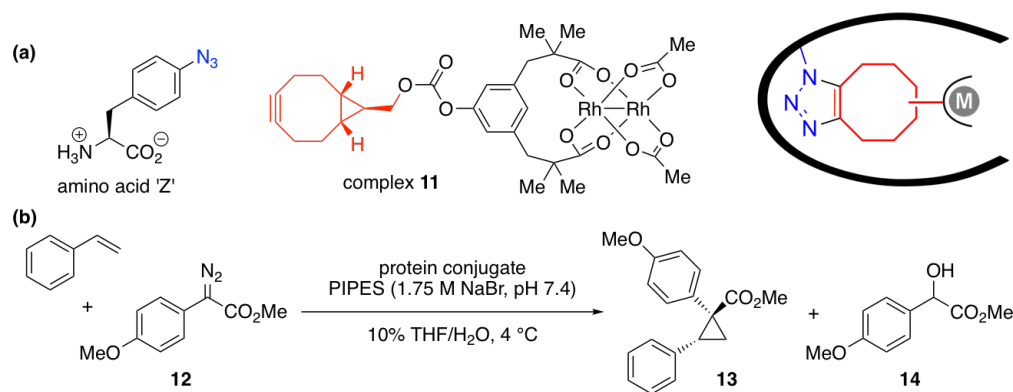


**Figure 1-8.** Examples of hybrid catalysts based on the biotin-streptavidin technology which have been optimized by directed evolution. (a) Biotin linker for attachment of organometallic catalysts. (b) Schematic of catalyst design. (c) Enantioselective



hydrogenation of alkenes with Rh-streptavidin metalloenzymes reported by Reetz and co-workers.<sup>56</sup> **(d)** Reduction of ketones to alcohols with Ru-streptavidin metalloenzymes reported by Ward and co-workers.<sup>57</sup> **(e)** Suzuki cross-coupling catalyzed by Pd-streptavidin metalloenzymes reported by Ward and co-workers.<sup>58</sup>

Lewis and co-workers have reported a design strategy that relies on the strain-promoted azide-alkyne cycloaddition to covalently attach a rhodium complex to the interior of a protein (Figure 1-9).<sup>60</sup> Proline oligopeptidase from *Pyrococcus furiosus* (*Pf*POP) was chosen due to its hyperthermostability and large interior cavity, which was expected to envelop Rh-complex **11**, thus enhancing the ability of the protein to influence activity. Rationally introduced mutations guided by a homology model led to the creation of a hybrid catalysts which perform cyclopropanation of styrene using diazo compound **12** with good activity and enantioselectivity. The challenge of creating and screening larger numbers of variants (hundreds) is in part due to the significant cyclopropanation activity of the free cofactor, which forms racemic product. As a result, creation of the hybrid catalyst is complicated by the need to remove unbound cofactor. An efficient and higher throughput procedure was recently developed for the assembly of hybrid *Pf*POP-Rh-complex catalysts with subsequent removal of unbound catalyst.<sup>61</sup> This procedure enabled directed evolution of the hybrid catalyst via random mutagenesis which ultimately identified an enzyme variant with superior cyclopropanation activity compared with the previous variant from the rational design effort.



**Figure 1-9.** A platform for Rh-protein conjugates based on proline peptidase.<sup>60,61</sup> **(a)** Strategy for formation of artificial metalloenzyme. **(b)** Carbene transfer catalyzed by enzyme variants.

## 1.7 Outlook

A key question for abiological catalysis with proteins is: what can the protein impart that cannot be readily achieved with small-molecule catalysts? We believe this is partially a substrate-binding problem. In addition to rate acceleration, such binding will enable precise control of transition-state orientation; this can impart selectivity for kinetically disfavored products, as in the cis-cyclopropanation of alkenes<sup>38</sup> or the amination of homobenzylic C–H bonds<sup>48</sup> catalyzed by engineered variants of P450<sub>BM3</sub>. Furthermore, the ability to improve enzyme activity through directed evolution may allow earth-abundant metals, like Fe and Co, to catalyze reactions that are typically associated with more precious transition metals, like Rh and Ir, provided a protein with trace levels of activity can be discovered. Careful analysis of side products from enzyme-catalyzed reactions may even be a feasible route for the invention of uncharted chemical transformations.

Synthetic chemistry has been revolutionized by the use of enzymes, which catalyze natural reactions with high activity and selectivity, often under mild conditions and without the need for protecting groups. We anticipate that directed evolution of artificial metalloenzymes will extend these advantages to non-natural reactions, including those currently inaccessible to small molecule catalysts.

## 1.8 References and Notes

---

1. He, J.; Muller, M.; Hertweck, C. Formation of the aureothin tetrahydrofuran ring by a bifunctional cytochrome P450 monooxygenase. *J. Am. Chem. Soc.* **2004**, *126*, 16742–16743.
2. Schmartz, P. C.; Wolfel, K.; Zerbe, K. et al. Substituent effects on the phenol coupling reaction catalyzed by the vancomycin biosynthetic P450 enzyme OxyB. *Angew. Chem. Int. Ed.* **2012**, *51*, 11468–11472.
3. Barry, S. M.; Kers, J. A.; Johnson E. G. et al. Cytochrome P450-catalyzed L-tryptophan nitration in thaxtomin phytotoxin biosynthesis. *Nat. Chem. Biol.* **2012**, *8*, 814–816.

4. McIntosh, J.; Farwell, C. C.; Arnold, F. H. Expanding P450 catalytic reaction space through evolution and engineering. *Curr. Opin. Chem. Biol.* **2014**, *19*, 126–134.
5. O'Brien, P. J.; Herschlag, D. Catalytic promiscuity and the evolution of new enzymatic activities. *Chem. Biol.* **1999**, *6*, R91–R105.
6. Copley, S. D. Enzymes with extra talents: moonlighting functions and catalytic promiscuity. *Curr. Opin. Chem. Biol.* **2003**, *7*, 265–272.
7. Khersonsky, O.; Roodveldt, C.; Tawfik, D. Enzyme promiscuity: evolutionary and mechanistic aspects. *Curr. Opin. Chem. Biol.* **2006**, *10*, 498–508.
8. Bloom, J. D.; Labthavikul, S. T.; Otey, C. R.; Arnold, F. H. Protein stability promotes evolvability. *Proc Natl Acad Sci U S A* **2006**, *103*, 5869–5874.
9. Hammer, S. C.; Kubik, G.; Watkins, E.; Huang, S.; Minges, H.; Arnold, F. H. Anti-Markovnikov alkene oxidation by metal-oxo-mediated enzyme catalysis. *Science* **2017**, *358*, 215–218.
10. Romney, D. K.; Murciano-Calles, J.; Wehrmuller, J. E.; Arnold, F. H. Unlocking reactivity of TrpB: a general biocatalytic platform for synthesis of tryptophan analogs. *J. Am. Chem. Soc.* **2017**, *139*, 10769–10776.
11. Bloom, J. D.; Arnold, F. H. In the light of directed evolution: Pathways of adaptive protein evolution. *Proc Natl Acad Sci U S A* **2009**, *106*, 9995–10000.
12. Morley, K. L.; Kazlauskas, R. J. Improving enzyme properties: when are closer mutations better? *Trends Biotechnol.* **2005**, *23*, 231–237.
13. Kille, S.; Acevedo-Rocha, C. G.; Parra, L. P.; Zhang, Z. G.; Opperman, D. J.; Reetz, M. T.; Acevedo, J. P. Reducing codon redundancy and screening effort of combinatorial protein libraries created by saturation mutagenesis. *ACS Synth. Biol.* **2013**, *2*, 83–92.
14. Stemmer, W. P. C. Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature* **1994**, *370*, 389–391.
15. Romero, P. A.; Arnold, F. H. Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 866–876.
16. Jackel, C.; Kast, P.; Hilvert, D. Protein design by directed evolution. *Annu Rev Biophys.* **2008**, *37*, 153–173.

- 
17. Packer, M. S.; Liu, D. R. Methods for the directed evolution of proteins. *Nat. Rev. Genet.* **2015**, *16*, 379–394.
  18. Smith J. M. Natural selection and the concept of a protein space. *Nature* **1970**, *225*, 563–564.
  19. Chu, S. Carbon capture and sequestration. *Science* **2009**, *325*, 1599.
  20. Alvizo, O.; Nguyen, L. J.; Savile, C. K. et al. Directed evolution of an ultrastable carbonic anhydrase for highly efficient carbon capture from flue gas. *Proc Natl Acad Sci U S A* **2014**, *111*, 16436–16441.
  21. O’Loughlin, T. L.; Patrick, W. M.; Matsumura, I. Natural history as a predictor of protein evolvability. *Protein Eng. Des. Sel.* **2006**, *19*, 439–442.
  22. Kille, S.; Zilly, F. E.; Acevedo, J. P.; Reetz, M. P. Regio- and stereoselectivity of P450-catalysed hydroxylation of steroids controlled by laboratory evolution. *Nat. Chem.* **2011**, *3*, 738–743.
  23. For a recent review on this topic, see: Roiban, G. D.; Reetz, M. T. Expanding the toolbox of organic chemists: directed evolution of P450 monooxygenases as catalysts in regio- and stereoselective oxidative hydroxylation. *Chem Commun.* **2015**, *51*, 2208–2224.
  24. Meinhold, P.; Peters, M. W.; Chen, M. Y. et al. Direct conversion of ethane to ethanol by engineered cytochrome P450 BM3. *ChemBioChem* **2005**, *6*, 1765–1768.
  25. Xu, F.; Bell, S. G.; Lednik, J. et al. The heme monooxygenase Cytochrome P450cam can be engineered to oxidize ethane to ethanol. *Angew. Chem. Int. Ed.* **2005**, *44*, 4029–4032.
  26. Fasan, R.; Chen, M. M.; Crook, N. C.; Arnold, F. H. Engineered alkane-hydroxylating cytochrome P450<sub>BM3</sub> exhibiting natively like catalytic properties. *Angew. Chem. Int. Ed.* **2007**, *46*, 8414–8418.
  27. Fasan, R.; Meharena, Y. T.; Snow, C. D.; Poulos, T. L.; Arnold, F. H. Evolutionary history of a specialized P450 propane monooxygenase. *J. Mol. Biol.* **2008**, *383*, 1069–1080.
  28. Zhang, K.; Damaty, S. E.; Fasan, R. P450 fingerprinting method for rapid discovery of terpene hydroxylation P450 catalysts with diversified regioselectivity. *J. Am. Chem. Soc.* **2011**, *133*, 3242–3245.
  29. Kolev, J. N.; O’Dwyer, K. M.; Jordan, C. T.; Fasan, R. Discovery of potent parthenolid-

- based antileukemic agents enabled by late-stage P450-mediated C–H functionalization. *ACS Chem. Biol.* **2014**, *9*, 164–173.
30. Neufeld, K.; Henßen, B.; Pietruszka, J. Enantioselective allylic hydroxylation of  $\omega$ -alkenoic acids and esters by P450 BM3 monooxygenase. *Angew. Chem. Int. Ed.* **2014**, *53*, 13253–13257.
31. For recent reviews on this topic, see: (a) Hyster, T. K.; Ward, T. R. Genetic optimization of metalloenzymes: enhancing enzymes for non-natural reactions. *Angew. Chem. Int. Ed.* **2016**, *55*, 7344–7357. (b) Prier, C. K.; Arnold, F. H. Chemomimetic biocatalysis: exploiting the synthetic potential of cofactor-dependent enzymes to create new catalysts. *J. Am. Chem. Soc.* **2015**, *137*, 13992–14006. (c) Renata H.; Wang, Z. J.; Arnold, F. H. Expanding the enzyme universe: accessing non-natural reactions by mechanism-guided directed evolution. *Angew. Chem. Int. Ed.* **2015**, *54*, 3351–3367.
32. Sreenilayam, G.; Moore, E. J.; Steck, V.; Fasan, R. Stereoselective olefin cyclopropanation under aerobic conditions with an artificial enzyme incorporating an iron-chlorin e6 cofactor. *ACS Catal.* **2017**, *7*, 7629–7633.
33. Reynolds, E. W.; McHenry, M. W.; Cannac, F.; Gober, J. G.; Snow, C. D.; Brustad, E. M. An evolved orthogonal enzyme/cofactor pair. *J. Am. Chem. Soc.*, **2016**, *138*, 12451–12458.
34. Mansuy, D. Cytochrome P-450 and synthetic models. *Pure & Appl Chem.* **1987**, *59*, 759–770.
35. Robbins Wolf, J.; Hamaker C. G.; Djukic J.-P.; Kodadek, T.; Woo, L. K. Shape and stereoselective cyclopropanation of alkenes catalyzed by iron porphyrins. *J. Am. Chem. Soc.* **1995**, *117*, 9194–9199.
36. Baumann, L. K.; Mbuvi, H. M.; Du, G.; Woo, L. K. Iron porphyrin catalyzed N–H insertion reactions with ethyl diazoacetate. *Organometallics* **2007**, *26*, 3995–4002.
37. Mirafzal, G. A.; Cheng, G.; Woo, K. L. A new and efficient method for the selective olefination of aldehydes with ethyl diazoacetate catalyzed by an iron(II) porphyrin complex. *J. Am. Chem. Soc.* **2002**, *124*, 176–177.
38. Coelho, P. S.; Brustad, E. M.; Kannan, A.; Arnold, F. H. Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes. *Science* **2013**, *339*, 307–310.

- 
39. Brandenburg, O. F.; Prier, C. K.; Chen, K.; Knight, A. M.; Wu, Z.; Arnold, F. H. Stereoselective enzymatic synthesis of heteroatom-substituted cyclopropanes. *ACS Catal.* **2018**, *8*, 2629–2634.
  40. Chen, K.; Huang, X.; Kan, S. B. J.; Zhang, R. K.; Arnold, F. H. Enzymatic construction of highly strained carbocycles. *Science* **2018**, *360*, 71–75.
  41. Wang, Z. J.; Peck, N. E.; Renata, H.; Arnold, F. H. Cytochrome P450-catalyzed insertion of carbenoids into N–H bonds. *Chem. Sci.* **2014**, *5*, 598–601.
  42. Bordeaux, M.; Tyagi, V.; Fasan, R. Highly diastereoselective and enantioselective olefin cyclopropanation using engineered myoglobin-based catalysts. *Angew. Chem. Int. Ed.* **2015**, *54*, 1744–1748.
  43. Kan, S. B. J.; Lewis, R. D.; Chen, K. & Arnold, F. H. Directed evolution of cytochrome *c* for carbon–silicon bond formation: Bringing silicon to life. *Science* **2016**, *354*, 1048–1051.
  44. Kan, S. B. J.; Huang, X.; Gumulya, Y.; Chen, K.; Arnold, F. H. Genetically programmed chiral organoborane synthesis. *Nature* **2017**, *552*, 132–136.
  45. Svastits E. W.; Dawson, J. H.; Breslow, R.; Gellman, S. H. Functionalized nitrogen atom transfer catalyzed by cytochrome P-450. *J. Am. Chem. Soc.* **1985**, *107*, 6427–6428.
  46. McIntosh J. A.; Coelho, P. S.; Farwell, C. C.; Wang, Z. J.; Lewis, J. C.; Brown, T. R.; Arnold, F. H. Enantioselective intramolecular C–H amination catalyzed by engineered cytochrome P450 enzymes in vitro and in vivo. *Angew. Chem. Int. Ed.* **2013**, *52*, 9309–9312.
  47. Singh, R.; Bordeaux, M.; Fasan, R. P450-catalyzed intramolecular *sp*<sup>3</sup> C–H amination with arylsulfonyl azide substrates. *ACS Catal.* **2014**, *4*, 546–552.
  48. Hyster, T. K.; Farwell, C. C.; Buller, A. R.; McIntosh, J. A.; Arnold, F. H. Enzyme-controlled nitrogen-atom transfer enables regiodivergent C–H amination. *J. Am. Chem. Soc.* **2014**, *136*, 15505–15508.
  49. Farwell, C. C. et al. Enantioselective imidation of sulfides via enzyme-catalyzed intermolecular nitrogen-atom transfer. *J. Am. Chem. Soc.* **2014**, *136*, 8766–8771.
  50. Cho, I.; Prier, C. K.; Jia, Z.-J.; Zhang, R. K.; Görbe, T.; Arnold, F. H. Enantioselective aminohydroxylation of styrenyl olefins catalyzed by an engineered hemoprotein. *Angew. Chem. Int. Ed.* **2019**, just accepted. DOI: 10.1002/anie.201812968.

- 
51. Key, H. M.; Dydio, P.; Clark, D. S.; Hartwig, J. F. Abiological catalysis by artificial haem proteins containing noble metals in place of iron. *Nature* **2016**, *534*, 534–537.
  52. Dydio, P.; Key, H. M.; Nazarenko, A.; Rha, J. Y.-E.; Seyedkazemi, V.; Clark, D. S.; Hartwig, J. F. An artificial metalloenzyme with the kinetics of native enzymes. *Science* **2016**, *354*, 102–106.
  53. Key, H. M.; Dydio, P.; Liu, Z.; Rha, J. Y.-E.; Nazarenko, A.; Seyedkazemi, V.; Clark, D. S.; Hartwig, J. F. Beyond iron: Iridium-containing P450 enzymes for selective cyclopropanations of structurally diverse alkenes. *ACS Cent. Sci.*, **2017**, *3*, 302–308.
  54. Dydio, P.; Key, H. M.; Hayashi, H.; Clark, D. S.; Hartwig, J. F. Chemoselective, enzymatic C–H bond amination catalyzed by a cytochrome P450 containing an Ir(Me)-PIX cofactor. *J. Am. Chem. Soc.* **2017**, *139*, 1750–1753.
  55. Song, W. J.; Tezcan, F. A. A designed supramolecular protein assembly with in vivo enzymatic activity. *Science* **2014**, *346*, 1525–1528.
  56. Reetz, M. T.; Peyralans, J. J.-P.; Maichele, A.; Fu, Y.; Maywald, M. Directed evolution of hybrid enzymes: Evolving enantioselectivity of an achiral Rh-complex anchored to a protein. *Chem. Commun.* **2006**, *41*, 4318–4320.
  57. Creus, M.; Pordea, A.; Rossel, T.; Sardo, A.; Letondor, C.; Ivanova, A.; LeTrong, I.; Stenkamp, R. E.; Ward, T. R. X-ray structure and designed evolution of an artificial transfer hydrogenase. *Angew. Chem. Int. Ed.* **2008**, *47*, 1400–1404.
  58. Chatterjee, A.; Mallin, H.; Klehr, J.; Vallapurackal, J.; Finke, A. D.; Vera, L.; Marsh M.; Ward, T. R. An enantioselective artificial Suzukiase based on the biotin–streptavidin technology. *Chem. Sci.* **2016**, *7*, 673–677.
  59. Jeschek, M.; Reuter, R.; Heinisch, T.; Trindler, C.; Klehr, J.; Panke, S.; Ward, T. R. Directed evolution of artificial metalloenzymes for in vivo metathesis. *Nature* **2016**, *537*, 661–665.
  60. Srivastava, P.; Yang, H.; Ellis-Guardiola, K.; Lewis, J. C. Engineering a dirhodium artificial metalloenzyme for selective olefin cyclopropanation. *Nat. Commun.* **2015**, *6*, DOI: 10.1038/ncomms8789.
  61. Yang, H. et al. Evolving artificial metalloenzymes via random mutagenesis. *Nat. Chem.* **2018**, *10*, 318–324.

## ENANTIOSELECTIVE ENZYME-CATALYZED AZIRIDINATION ENABLED BY EVOLUTION OF A CYTOCHROME P450

Material for this chapter appears in Farwell, C. C.<sup>†</sup>; **Zhang, R. K.**<sup>†</sup>; McIntosh, J. A.; Hyster, T. K.; Arnold F. H. “Enantioselective enzyme-catalyzed aziridination enabled by active-site evolution of a cytochrome P450,” *ACS Central Science* **2015**, *1*, 89–93. **DOI:** 10.1021/acscentsci.5b00056. (<sup>†</sup>Denotes equal contribution) This work was performed in collaboration with all authors. Reprinted with permission from American Chemical Society.

### 2.1 Abstract

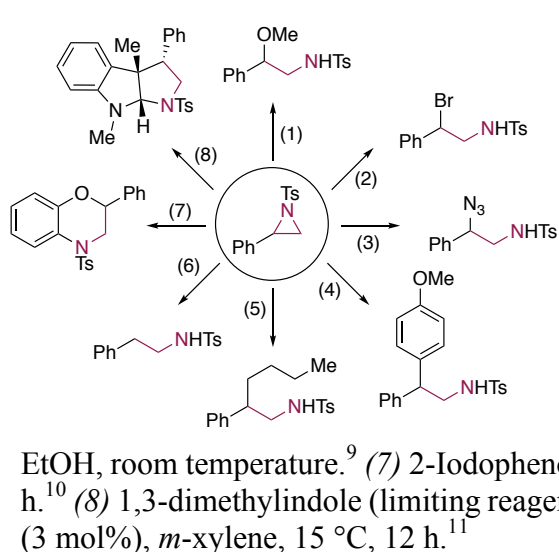
One of the greatest challenges in protein design is creating new enzymes. But nature does this all the time, in response to new challenges and opportunities. Borrowing from nature's evolutionary strategy, we engineered a bacterial cytochrome P450 through a process of accumulating beneficial mutations to catalyze highly enantioselective olefin aziridination, a synthetically useful reaction. The new enzyme is genetically encoded, functions *in vitro* or in whole cells, and has been engineered to exhibit high selectivity (up to 99% ee) and productivity (up to 1,000 total turnovers) for intermolecular aziridination, demonstrated here with tosyl azide and substituted styrenes.

### 2.2 Introduction

Aziridines are three-membered saturated nitrogen-containing heterocycles. This motif has attracted significant interest due to its utility as a building block in chemical synthesis and its presence in biologically active molecules.<sup>1,2</sup> The inherent ring strain of this heterocycle coupled with the electrophilicity of the nitrogen atom, which may be modulated by attached groups, allows this motif to undergo ring cleavage and ring expansion reactions. Many useful motifs, including the ones shown in Figure 2-1, can be accessed from the aziridine. In addition, certain functionalized aziridines, such as vinylaziridines and ethynylaziridines, can participate in rearrangement and additional

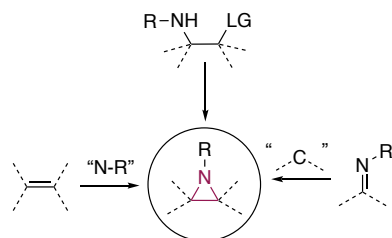


cycloaddition reactions.<sup>3</sup> Indeed, due to the reactivity of the aziridine ring, bioactive molecules which contain this group often rely on its reactivity for their mechanism of action.<sup>2</sup>



**Figure 2-1.** Select examples of reactions of aziridines, using 2-phenyl-*N*-tosylaziridine as a model. (1) MeOH (as solvent), H<sub>2</sub>SO<sub>4</sub> (3 equiv), continuous flow conditions.<sup>4</sup> (2) *n*-Bu<sub>4</sub>N<sup>+</sup> Br<sup>-</sup>, BF<sub>3</sub>•OEt<sub>2</sub>, CH<sub>2</sub>Cl<sub>2</sub>, 0°C, 2 min.<sup>5</sup> (3) NaN<sub>3</sub>, ammonium-12-molybdophosphate (10 mol%), MeCN, room temperature.<sup>6</sup> (4) Anisole (2 equiv), [Ag(COD)<sub>2</sub>]PF<sub>6</sub> (2 mol%), DCE, room temperature, 3 h.<sup>7</sup> (5) *n*-BuZnBr (3 equiv), NiCl<sub>2</sub>•glyme (5 mol%), dimethyl fumarate (10 mol%), dioxane/DMA, 23 °C, 6 h.<sup>8</sup> (6) Polymethylhydrosiloxane, Pd/C, EtOH, room temperature.<sup>9</sup> (7) 2-Iodophenol, Cu/Al<sub>2</sub>O<sub>3</sub> (4 mol%), K<sub>2</sub>CO<sub>3</sub>, DMF, 100 °C, 8 h.<sup>10</sup> (8) 1,3-dimethylindole (limiting reagent), [(MeCN)<sub>4</sub>Cu]BF<sub>4</sub> (5 mol%), (*R*)-XylBINAP (3 mol%), *m*-xylene, 15 °C, 12 h.<sup>11</sup>

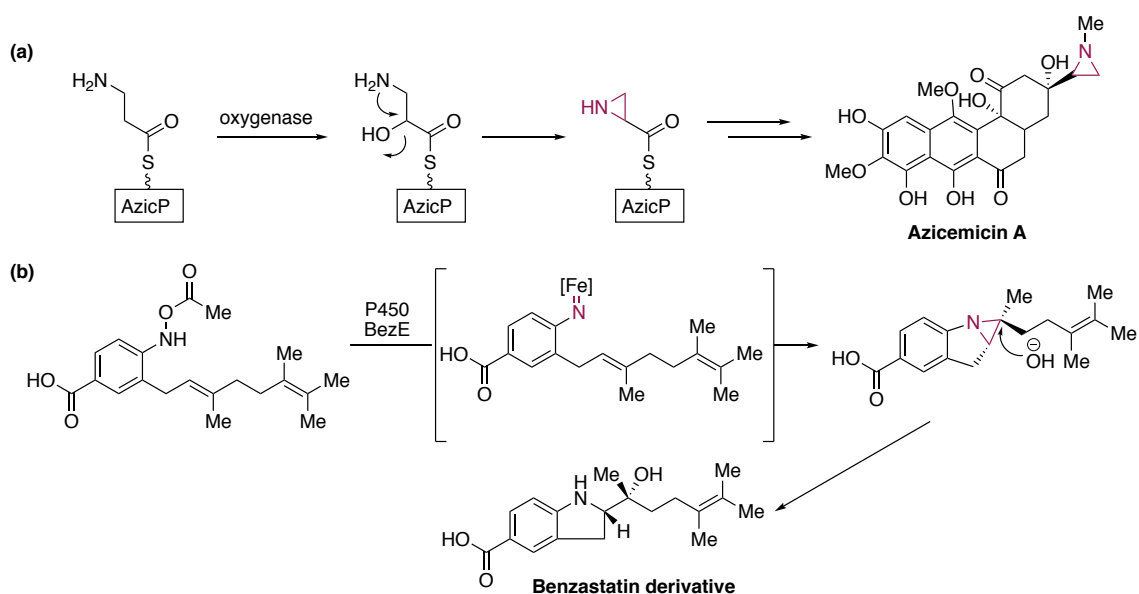
Several strategies have been developed to access aziridines.<sup>1</sup> The primary routes include delivery of a suitable nitrogen source to alkenes, addition of a carbon source to imines, and intramolecular cyclization of an amine with an appropriately placed leaving group (Figure 2-2). Of these approaches, we found the transformation of alkenes with a suitable nitrogen source to be particularly attractive, as alkenes are prevalent functional groups and typically stable under aqueous conditions. Aziridination of alkenes using reagents such as azides,<sup>12–15</sup> iminoiodinanes,<sup>16</sup> and haloamines<sup>17</sup> is known to be catalyzed by transition metal catalysts, including metalloporphyrins, and occurs presumably through a metal-nitrene intermediate.



**Figure 2-2.** Strategies to access aziridines.<sup>1</sup>

By contrast, enzymes which perform nitrogen-atom transfer chemistry through a

metal-nitrene intermediate are extremely rare. As a result, the biosynthesis of naturally occurring aziridine rings, for instance that in azicemicin,<sup>18</sup> is believed to typically proceed by intramolecular nucleophilic displacement of an appropriately placed leaving group (Figure 2-3a). Notably, after the publication of the work presented in this chapter, Ohnishi and co-workers discovered a cytochrome P450 from *Streptomyces* sp. R118, BezE, involved in the biosynthesis of benzastatins, which may perform aziridination through nitrene transfer as its native function (Figure 2-3b).<sup>19</sup> This discovery demonstrates how biocatalysis can inform mechanistic possibilities for enzymes in complex biosynthetic pathways.



**Figure 2-3.** Natural biosynthetic routes to aziridines. **(a)** Proposed strategy for aziridine formation in azicemicins.<sup>18</sup> **(b)** Proposed nitrene transfer mechanism of BezE, a cytochrome P450 enzyme involved in benzastatin biosynthesis.<sup>19</sup>

In 2014, our group reported sulfide imidation with tosyl azide catalyzed by variants of cytochrome P450<sub>BM3</sub> from *Bacillus megaterium*.<sup>20</sup> In this work, it was observed that the reaction was promoted by electron-rich sulfides. Given the less nucleophilic nature of alkenes relative to sulfides, we surmised that intermolecular aziridination would be a more challenging activation mode for the enzyme. We hypothesized, however, that protein engineering would allow us to circumvent these limitations if the enzyme could more

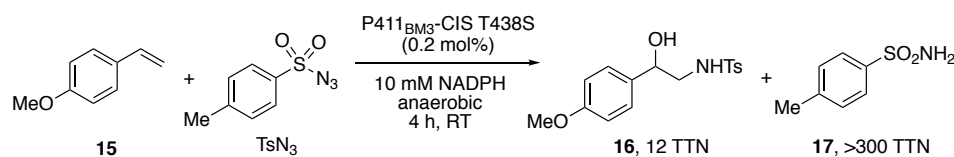
effectively bind and orient the substrates in the enzyme active site for productive nitrene transfer. Here we demonstrate that active site engineering can indeed enable a heme protein to catalyze efficient, highly enantioselective intermolecular nitrene transfer to alkenes to make aziridines.

## 2.3 Results and discussion

### 2.3.1 Identification of heme proteins with promiscuous aziridination activity

We started this investigation of enzyme-catalyzed aziridination with an engineered variant of cytochrome P450<sub>BM3</sub>, P411<sub>BM3</sub>-CIS T438S, which differs from wild type by 15 mutations (see Table A-1 in Appendix A for mutations from wild type P450<sub>BM3</sub>). This variant was previously found to be effective for intramolecular C–H amination<sup>21</sup> and sulfide imidation.<sup>20</sup> We call this enzyme a “P411” due to the change in the characteristic CO-bound Soret peak from 450 to 411 nm effected by mutation of the cysteine residue that coordinates the heme iron to serine (C400S).<sup>22</sup> This axial cysteine is completely conserved in cytochrome P450s and is required for the native monooxygenase activity.<sup>23,24</sup> Thus, the P411 enzyme is no longer a “cytochrome P450”. Two crystal structures of P411 variants of P450<sub>BM3</sub> show that S400 coordinates the iron and causes no significant structural perturbation in the substrate binding pocket.<sup>22,25</sup>

Previous work on enzymatic sulfide imidation suggested that electron-rich sulfides promote nitrene transfer.<sup>20</sup> Reasoning that the electronic properties of the alkene substrate could influence aziridination as well, we tested the activity of P411<sub>BM3</sub>-CIS T438S toward the electron-rich 4-methoxystyrene using tosyl azide (TsN<sub>3</sub>) as the nitrene precursor (Figure 2-4). Tosyl azide was completely consumed in this reaction, the major product of which was the azide reduction product *p*-toluenesulfonamide (**17**, >300 total turnovers (TTN), not shown in Table 1). Amidoalcohol **16** appeared as a minor product. Control experiments showed that the desired aziridine product rapidly decomposes under aqueous reaction conditions to the corresponding amidoalcohol **16** (Figure A-1 in Appendix A). Degradation of this aziridine product has also been observed in studies with small-molecule catalysts.<sup>26</sup>



**Figure 2-4.** Initial olefin aziridination activity catalyzed by purified holoenzyme P411<sub>BM3</sub>-CIS T438S. Reactions were performed in 0.1 M KPi buffer pH = 8.0, with 7.5 mM olefin and 2.5 mM tosyl azide. TTN, total turnover number; TTNs were determined by HPLC analysis. KPi, potassium phosphate.

We therefore inferred that production of **16** is directly related to the nitrene transfer activity of the enzyme toward olefin **15**. The low level of nitrene transfer activity to 4-methoxystyrene using P411<sub>BM3</sub>-CIS T438S prompted us to investigate other heme proteins for this activity. Myoglobin (horse heart), cytochrome *c* (bovine heart), cytochrome P450<sub>Rhf</sub> (from *Rhodococcus* sp. NCIMB 9784), and the heme cofactor alone (iron protoporphyrin IX) were all inactive for this intermolecular aziridination (Table A-2 in Appendix A). An engineered variant of the thermostable cytochrome P450 from *Sulfolobus acidocaldarius*, CYP119, that contained an axial cysteine-to-serine mutation (C317S) did catalyze low levels of aziridination (~ 7 TTN). This demonstrates that mutations previously described to activate non-natural nitrene-transfer activity in P450<sub>BM3</sub> may confer measurable activity on other P450s as well. In turn, these enzymes should be suitable starting points for further engineering.<sup>27</sup>

### 2.3.2 Active-site engineering of an enzyme aziridination catalyst

We also tested a set of P450<sub>BM3</sub> variants (Table A-3 in Appendix A). Proteins lacking the C400S and/or T268A mutations, mutants of P411<sub>BM3</sub>-CIS T438S with different axial mutations, and other enzymes differing from P411<sub>BM3</sub>-CIS T438S by 2–5 alanine mutations in the active site showed low or no aziridination activity (< 8 TTN). Of all the enzymes tested, a variant of P411<sub>BM3</sub>-CIS T438S having a single active site substitution, I263F, was the most active toward 4-methoxystyrene by a wide margin, providing 150 total turnovers in the formation of amidoalcohol **16** from 4-methoxystyrene. This variant, which was found to promote regioselective intramolecular C–H amination in a previous study,<sup>25</sup> supported aziridination at 10-fold increased total turnovers compared

to the P411<sub>BM3</sub>-CIS T438S parent enzyme (henceforth referred to as the P enzyme). Additionally, this I263F-containing variant was active for aziridination on less electron-rich olefins 4-methylstyrene (**18**) and styrene (**20**), for which the corresponding aziridine products **19** and **21** were stable under the reaction conditions (Table 2-1). Hereafter P411<sub>BM3</sub>-CIS T438S will be referred to as P (for parent) and by extension P411<sub>BM3</sub>-CIS T438S I263F will be called P-I263F.

**Table 2-1.** Total turnovers (TTN) to product for aziridination catalyzed by purified holoenzymes P411<sub>BM3</sub>-CIS T438S (P) and P411<sub>BM3</sub>-CIS T438S I263F (P-I263F) with selected styrenes and tosyl azide.<sup>a</sup>

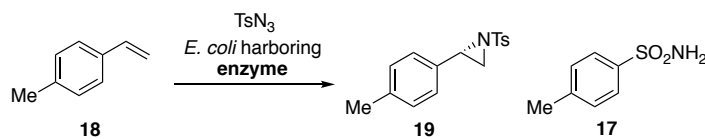
<p> <math>\text{R} = \text{OMe}, \mathbf{15}</math>  <math>\text{R} = \text{Me}, \mathbf{18}</math>  <math>\text{R} = \text{H}, \mathbf{20}</math> </p> <p> <math>\text{R} = \text{Me}, \mathbf{19}</math>  <math>\text{R} = \text{H}, \mathbf{21}</math> </p>			
Enzyme	TTN <b>16</b>	TTN <b>19</b>	TTN <b>21</b>
P411 <sub>BM3</sub> -CIS T438S (P)	15	8	5
P-I263F	150	160	190

<sup>a</sup>Reactions were performed in 0.1 M KPi buffer pH = 8.0 with 7.5 mM olefin and 2.5 mM tosyl azide, using purified holoproteins.

The aziridination reaction could also be carried out using whole *Escherichia coli* cells expressing the P-I263F enzyme. No aziridine product was observed when cells not expressing the P411 catalyst were used, although tosyl azide was converted to sulfonamide **17** over the course of the reaction (Table 2-2). Reactions with *E. coli* cells harboring the P-I263F enzyme provided enough aziridine product to allow for screening variants in 96-well plate format. Thus, we sought further improvement by mutagenesis and screening for aziridination productivity. Site-saturation mutagenesis (SSM) libraries were created at several active site positions that were previously shown to influence productivity and enantioselectivity in other non-natural reactions (A78, L181, T438, A328).<sup>25,28</sup> Screening of these single SSM libraries for aziridination of 4-methylstyrene (**19**) identified P-I263F A328V, which performed the reaction with slightly improved yield and substantially improved %ee (96% ee; entry 4, Table 2-2). Another round of SSM performed on this

variant at additional active site positions (F87, T268, L437) resulted in P-I263F A328V L437V with improved aziridine yield and a further increase in enantioselectivity (99% ee (*S*); entry 6, Table 2-2). The P-I263F L437V and P-I263F A328V mutants were both less selective than P-I263F A328V L437V, demonstrating that both new mutations contribute to the very high enantioselectivity.

**Table 2-2.** Improvement in yield and %ee for aziridine product **19** with active-site evolution of P411<sub>BM3</sub>CIS T438S (P).<sup>a</sup>

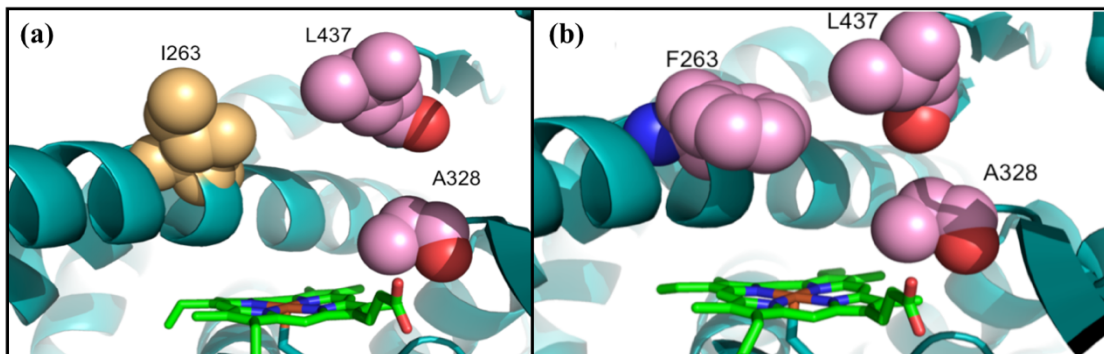


Entry	Enzyme	%yield <b>19</b>	%yield <b>17</b>	%ee <b>19</b> <sup>b</sup>
1	No enzyme <sup>c</sup>	0	95	N.A.
2	P411 <sub>BM3</sub> -CIS-T438S	1.1	95	25
3	P-I263F	40	54	58
4	P-I263F-A328V	43	50	96
5	P-I263F-L437V	37	52	92
6	<b>P-I263F-A328V-L437V</b>	<b>55</b>	<b>43</b>	<b>99</b>

<sup>a</sup> Reactions were carried out using whole *E. coli* cells resuspended in M9-N reaction buffer under anaerobic conditions, using 7.5 mM 4-methylstyrene and 2.5 mM tosyl azide. Yield is based on tosyl azide. See Experimental Methods (Section 2.5) for detailed reaction set up and quantification procedures. <sup>b</sup> % ee determined by SFC analysis and calculated as (*S* – *R*)/(*S* + *R*). <sup>c</sup> ‘No enzyme’ reactions were carried out using whole cells with no P411 enzyme expressed, as described in the experimental methods. N.A., not applicable.

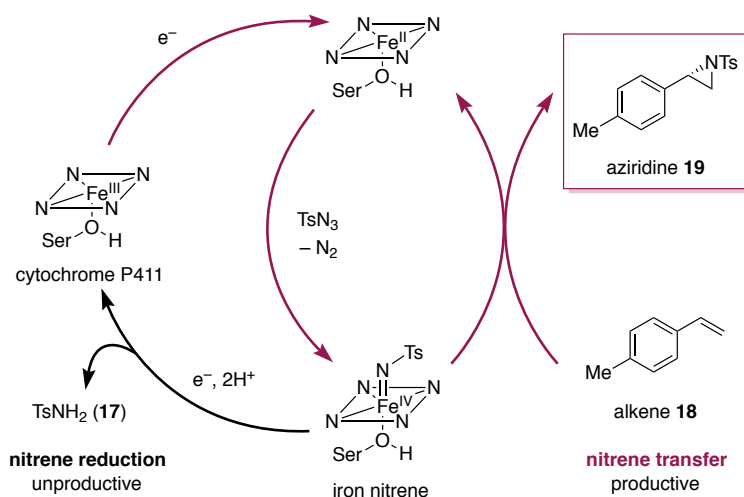
The highly enantioselective P-I263F A328V L437V variant has three mutations in its active site relative to the P enzyme used in initial reaction characterization (P411<sub>BM3</sub>-CIS T438S). The crystal structure of the heme domain of P-I263F was recently solved and shows how the F263 side chain is oriented toward the heme cofactor within the active site (PDB ID: 4WG2, Figure 2-5).<sup>25</sup> The effect of the I263F mutation on the active site is significant: the F263 side chain fills space above the heme cofactor whereas the I263 side chain is pointed up and away from the heme (Figure 2-5). Given the more conservative nature of the A328V and L437V mutations, these residues likely exert more subtle influences on active site structure, yet their impact on enantioselectivity is substantial (55%

ee with P-I263F vs 99% ee with P-I263F A328V L437V, using 4-methylstyrene, Table 2-2).



**Figure 2-5.** Comparison of active site residues in P411<sub>BM3</sub>-CIS and P-I263F. **(a)** P411<sub>BM3</sub>-CIS structure (PDB ID: 4H23) with I263 shown as van Der Waals spheres in gold, L437 and A328 shown in pink. The heme domain of P411<sub>BM3</sub>-CIS T438S is has not been solved, however this variant is only one mutation (T438S) different from P411<sub>BM3</sub>-CIS (shown here). **(b)** P-I263F structure (PDB ID: 4WG2) showing the active site and residues F263, A328 and L437 in pink.

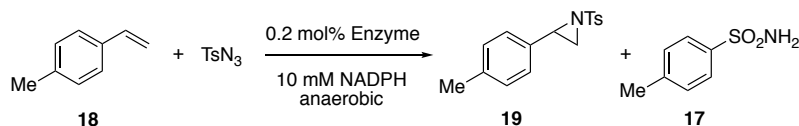
We proposed a hypothesis for the mechanism of this reaction, which is summarized in Figure 2-6. The reaction begins with the reduction of the Fe(III)-heme through gaining an electron from NADPH via the reductase domain. Control experiments show that aziridine **19** is not formed in the absence of NADPH when full-length P-I263F is used as catalyst in purified form (Table A-4 in Appendix A). Furthermore, no aziridine is formed when the heme domain of P-I263F is used as catalyst with NADPH as the reductant (Table A-4 in Appendix A), since NADPH is not capable of reducing P-I263F in the absence of the reductase domain (Figure A-2 in Appendix A). We propose that reaction of tosyl azide with ferrous heme results in the formation of an iron-nitrene species in the active site (formally in the +4 oxidation) which can be either intercepted by olefin to produce aziridine or reduced by a second electron transfer to form sulfonamide and return the catalyst to its ferric state.<sup>29</sup> Control experiments which monitored the formation of sulfonamide **17** under purified protein reaction conditions suggest that its formation is enzyme catalyzed (Table A-4 in Appendix A). In the event that sulfonamide is formed, the Fe(III)-heme must then consume another reducing equivalent to return to the reactive ferrous state.



**Figure 2-6.** Proposed mechanism of a cytochrome P411-catalyzed aziridination reaction. Reaction of tosyl azide with the ferrous heme cofactor generates an enzyme-bound iron nitrene intermediate. This nitrene then reacts with an olefin substrate, delivering an aziridine product. The nitrogen atoms in a plane represent the enzyme's heme cofactor. Ts = 4-toluenesulfonyl; Ser = serine. Notes: to date, the protonation state of the axial serine ligand in P411<sub>BM3</sub> enzymes is not known; a formal oxidation state of the iron nitrene intermediate is shown, but may not accurately reflect its electronic structure.

The improved aziridine yield could result from either an increase in the rate of aziridine formation or a decrease in the rate of competing azide reduction, or from a combination of both. To address this, we measured initial rates of reaction with the P-I263F, P-I263F A328V, and P-I263F A328V L437V enzymes as purified holoenzymes (Table 2-3, Figure A-3 in Appendix A). Initial rates of aziridination for the purified enzymes reflected the yield improvements observed in whole cells: P-I263F and P-I263F A328V have similar turnover frequencies (15–16 min<sup>-1</sup>), while P-I263F A328V L437V, having both new mutations, was improved (TOF ~24 min<sup>-1</sup>). The initial turnover frequency of sulfonamide formation *in vitro* was similar for all the enzymes, and faster than aziridine formation (TOFs ~26–29 min<sup>-1</sup>).



**Table 2-3.** Initial rates of aziridination and azide reduction for engineered enzymes.<sup>a</sup>

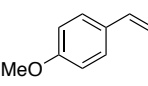
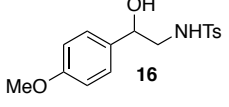
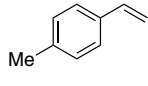
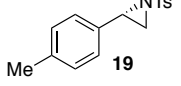
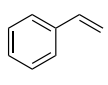
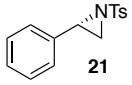
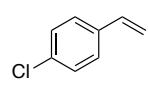
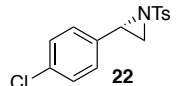
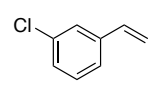
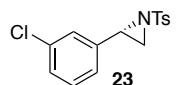
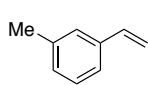
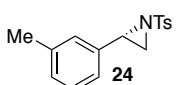
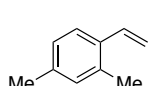
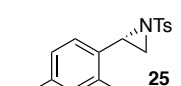
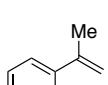
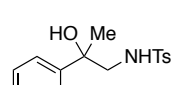
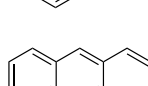
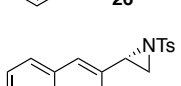
Enzyme	TOF 19 (min <sup>-1</sup> )	TOF 17 (min <sup>-1</sup> )	TOF 19/ TOF 17	TTN 19	TTN 17	TTN 19/ TTN 17
<b>P-I263F</b>	15	29	0.51	150	280	0.52
<b>P-I263F-A328V</b>	16	26	0.62	145	290	0.50
<b>P-I263F-A328V-L437V</b>	24	29	0.83	185	250	0.73

<sup>a</sup> Initial rates were measured using purified P411 enzymes with 7.5 mM olefin and 2.5 mM TsN<sub>3</sub>, as described in the Experimental Methods section (Section 2.5). Total turnover (TTN) values were determined using the same method as described for initial rates, with the exception that reactions were allowed to proceed for 4 hours in the anaerobic chamber. Data used to determine the initial rates is presented as Figure A-3. TOF, turnover frequency.

### 2.3.3 Substrate scope

Having obtained a highly active and selective enzyme variant, P-I263F A328V L437V, we investigated its substrate scope in whole cells using variously substituted styrenes and tosyl azide (Table 2-4). In addition to **15**, **18**, and **20**, P-I263F A328V L437V accepted electron deficient styrenes, differently substituted styrenes, and bulky vinylnaphthalene. Similar to 4-methoxystyrene,  $\alpha$ -methylstyrene gave exclusively racemic amido-alcohol hydrolysis product. We reason that this is likely the result of carbocation stabilization at the benzylic position afforded by resonance and hyperconjugative stabilization by the *p*-OMe and  $\alpha$ -Me groups respectively. The aziridine products were all highly enantioenriched though (Table 2-4). We demonstrated that with higher substrate concentration, the enzyme could catalyze higher turnovers to the aziridine product. For example, using 2-fold more styrene (15 mM) and 3-fold more tosyl azide (7.5 mM) compared to typical reaction conditions, the enzyme catalyzed 1,000 turnovers for aziridine **21**.

**Table 2-4.** Substrate scope of aziridination with P-I263F A328V L437V.<sup>a</sup>

Entry	Olefin	Product	TTN	% yield	% ee
1		 <b>16</b>	390	47	rac
2		 <b>19</b>	450 (960) <sup>b</sup>	55 (38) <sup>b</sup>	99 (99) <sup>b</sup>
3		 <b>21</b>	600 (1000) <sup>b</sup>	70 (40) <sup>b</sup>	99 (99) <sup>b</sup>
4		 <b>22</b>	290	36	99
5		 <b>23</b>	21	2	95
6		 <b>24</b>	85	10	95
7		 <b>25</b>	130	15	81
8		 <b>26</b>	83	10	rac
9		 <b>27</b>	53	6	88

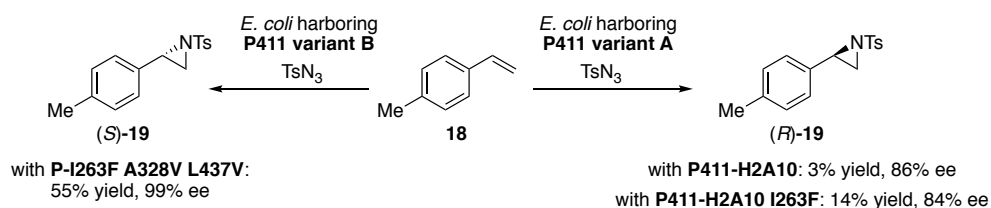
<sup>a</sup>Unless otherwise indicated, reactions were carried out using *E. coli* cells expressing P-I263F A328V L437V under anaerobic conditions, with 7.5 mM olefin and 2.5 mM tosyl azide. % ee determined as (S – R)/(S + R); absolute configurations were assigned based on analogy to **21**. Rac, racemic. See Experimental Methods (Section 2.5) for detailed procedures. <sup>b</sup>Reactions were carried out with 15 mM olefin and 7.5 mM tosyl azide.

Compared to small molecule transition metal catalysts, P-I263F A328V L437V catalyzes the olefin aziridination reaction with superior enantioselectivity. For example, the reaction of styrene with tosyl azide catalyzed by a chiral Ru-salen-CO complex affords aziridine **21** at 93% yield (982 turnovers) and 86% ee.<sup>1</sup> Using a chiral Co-porphyrin complex, aziridination of styrene with related nosyl azide affords the nosyl-protected aziridine in 75% yield (38 TTN) and 88% ee.<sup>30</sup> Comparatively, *E. coli* harboring P-I263F

A328V L437V delivers aziridine **21** with 99% ee (70% yield, 600 TTN) and related aziridines with good enantioselectivity (>80% ee). Notably though, enzymatic aziridination is performed at dilute substrate concentrations (significantly more dilute than with typical small molecule transition metal systems). Increasing the substrate concentration in enzyme-catalyzed reactions results in decreased yields, though the high enantioselectivity is maintained (entries 2 and 3, Table 2-4). This highlights a need for further investigation and optimization of these nitrene transfer enzymes.

### 2.3.4 Tuning the enantioselectivity of enzyme-catalyzed olefin aziridination

Either enantiomers of the aziridine product can be obtained by changing the enzyme catalyst. A different P411 variant, P411-H2A10, was found to perform the reaction between 4-methylstyrene and tosyl azide with the opposite stereochemical preference compared with P-I263F A328V L437V (Figure 2-7). Variant P411-H2A10 (15 amino acids from wild type), differs at six amino acid residues compared with P-I263F A328V L437V. Adding the I263F mutation to P411-H2A10 improved the yield without affecting enantioselectivity (Figure 2-7). This demonstrates the versatility of the P411 protein scaffold; the stereochemical preference of a reaction can be altered by introducing a small number of amino acid changes. The influence of protein sequence on product enantioselectivity has previously been demonstrated with P411-catalyzed sulfide imidation<sup>20</sup> and with other enzymes including carbonyl reductase<sup>31</sup> and halohydrin dehalogenase.<sup>32</sup> We did not pursue further efforts in improving the activity and enantioselectivity of P411-H2A10 I263F.



**Figure 2-7.** Enzyme-controlled enantioselectivity of olefin aziridination. Reactions were carried out using *E. coli* cells expressing the enzyme under anaerobic conditions, with 7.5 mM olefin and 2.5 mM tosyl azide. Absolute configurations were assigned based on analogy to **21**.

## 2.4 Conclusion

We report the first example of enzyme-catalyzed olefin aziridination. This challenging intermolecular reaction is catalyzed by a serine-ligated “P411” variant of cytochrome P450BM3. Mutations in the enzyme active site resulted in a variant that exhibits significantly improved azide utilization compared to the parental enzyme and high enantioselectivity (up to 99% ee). These results demonstrate the critical role of protein engineering in optimizing non-native activity and suggest that the well-known plasticity of the P450 active site can be leveraged to target synthetically useful reactions.

Findings from biocatalysis can additionally inform mechanistic possibilities for natural enzymes. Previously thought to be absent from natural enzyme mechanisms, after the publication of the work presented in this chapter, olefin aziridination by nitrene transfer was put forth as a possible natural mechanism for a P450 involved in benzastatin biosynthesis.<sup>19</sup>

This new aziridination biocatalyst is likely just one of many new catalysts that will be discovered when researchers start systematically exploring the new functions that existing enzymes can take on.<sup>33,34</sup> Exploiting the catalytic promiscuity of natural enzymes combined with evolutionary optimization will enable us to greatly expand the reaction space of genetically encoded and tunable biocatalysts.

## 2.5 Experimental Methods

See Appendix A for supporting tables and figures, characterization of compounds, details regarding calibration curves, assignment of absolute stereochemistry, and determination of enantioselectivity. HPLC calibration curves and NMR spectra are in the Supporting Information of the published paper, but not included here.

### 2.5.1 General information

Unless otherwise noted, all chemicals and reagents were obtained from commercial suppliers (Sigma-Aldrich, VWR, Alfa Aesar) and used without further purification. Silica

gel chromatography purifications were carried out using AMD Silica Gel 60, 230-400 mesh.  $^1\text{H}$  spectra were recorded on a Varian Inova 300 MHz or 500 MHz, or Bruker Prodigy 400 MHz instrument in  $\text{CDCl}_3$ , and are referenced to the residual solvent peak. Synthetic reactions were monitored using thin layer chromatography (Merck 60 gel plates) using an UV-lamp for visualization.

### 2.5.2 Chromatography

Analytical high-performance liquid chromatography (HPLC) was carried out using an Agilent 1200 series, and a Kromasil 100 C18 column (4.6 x 50 mm, 5  $\mu\text{m}$ ) with water and acetonitrile as the mobile phase. Semipreparative HPLC was performed using an Agilent XDB-C18 column (9.4 x 250 mm, 5  $\mu\text{m}$ ). Liquid-chromatography-mass-spectrometry (LC-MS) was carried out using an Agilent 6140 series equipped with a C18 column with water (+0.1% acetic acid) and acetonitrile as mobile phases. Analytical chiral HPLC was conducted using a supercritical fluid chromatography (SFC) system with isopropanol and liquid  $\text{CO}_2$  as the mobile phase. Chiral OB-H and AS-H columns were used to separate aziridine and amido-alcohol enantiomers (4.6 x 150 mm, 5  $\mu\text{m}$ ). Olefins were all commercially available; amido-alcohol<sup>35</sup> and aziridine<sup>36</sup> standards were prepared following literature procedures. %ee was calculated by dividing the major peak area by the sum of the peak areas determined by SFC chromatography.

### 2.5.3 Cloning and site-directed mutagenesis

pET22b(+) was used as a cloning and expression vector for all enzymes described in this study. P450 and P411 enzymes described in this study were expressed with a C-terminal 6xHis-tag. Site-directed mutagenesis was performed using a modified QuikChange<sup>TM</sup> mutagenesis protocol.<sup>37</sup> The PCR products were gel purified, digested with DpnI, repaired using Gibson Mix<sup>TM</sup>,<sup>38</sup> and directly transformed into *E. coli* strain BL21 *E. coli* cells (Lucigen).

### 2.5.4 Determination of P450 concentration

Concentration of P450 and P411 enzymes in whole cell experiments was determined from ferrous carbon monoxide binding difference spectra using previously reported extinction coefficients for cysteine-ligated ( $\epsilon = 91,000 \text{ M}^{-1} \text{ cm}^{-1}$ )<sup>39</sup> and serine-ligated enzymes ( $\epsilon = 103,000 \text{ M}^{-1} \text{ cm}^{-1}$ ).<sup>40</sup> When purified enzymes were used, concentration of P450 and P411 enzymes was accomplished by quantifying the amount of free hemin present in purified protein using the pyridine/hemochrome assay and the published extinction coefficient ( $\epsilon = 191,500 \text{ M}^{-1} \text{ cm}^{-1}$ ).<sup>41</sup>

### 2.5.5 Protein expression and purification

*E. coli* BL21 *E. coli* cells carrying a plasmid encoding a P450 or P411 protein were grown overnight in 25 mL Luria-Bertain medium with 0.1 mg/mL ampicillin (LB<sub>amp</sub>, 37 °C, 250 rpm). Hyperbroth medium (supplemented with glucose nutrient mix according to package instructions) with 0.1 mg/mL ampicillin (HB<sub>amp</sub>, 630 mL) in a 1-L flask was inoculated with 25 mL of the preculture and incubated at 37 °C and 230 rpm for 2.5 h. Cultures were then cooled on ice (30 min) and induced with 0.5 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) and 1.0 mM 5-aminolevulinic acid (final concentrations). Expression was conducted at 23 °C, 130 rpm, for 16–20 h. Cultures were then centrifuged (5,000 x g, 8 min, 4 °C) and the cell pellets frozen at –20 °C.

For purification, frozen cells were resuspended in buffer A (20 mM tris, 20 mM imidazole, 100 mM NaCl, pH 7.5, 4 mL/g of cell wet weight), loaded with 300  $\mu$ g/ml hemin, and disrupted by sonication (2 x 1 min, output control 5, 50% duty cycle; Sonicator 3000, Misonix, Inc.). To pellet insoluble material, lysates were centrifuged (20,000 x g for 0.5 h at 4 °C). P450 and P411 enzymes were expressed in a construct containing a C-terminal 6xHis-tag and were consequently purified using a nickel NTA column (1 mL HisTrap HP, GE Healthcare, Piscataway, NJ) using an AKTAexpress purifier FPLC system (GE healthcare). P450 or P411 enzymes were then eluted on a linear gradient from 0% buffer B (20 mM tris, 300 mM imidazole, 100 mM NaCl, pH 7.5) to 100 % buffer B over 10 column volumes (P450/P411 enzymes elute at around 80 mM imidazole). Fractions containing P450 or P411 enzymes were pooled, concentrated, and subjected to three

exchanges of phosphate buffer (0.1 M potassium phosphate (KPi), pH = 8.0) to remove excess salt and imidazole. Concentrated proteins were aliquoted, flash-frozen on powdered dry ice, and stored at -20 °C until later use.

#### **2.5.6 Typical procedure for small-scale aziridination bioconversions under anaerobic conditions using purified enzymes.**

Small-scale reactions (400  $\mu$ L) were conducted in 2 mL crimp vials (Agilent Technologies, San Diego, CA) containing buffer (0.1 M KPi, pH = 8.0), purified enzyme, and oxygen depletion system (GOX, 20X stock solution containing 14,000 U/mL catalase, 1,000 U/mL glucose oxidase dissolved in 0.1 M KPi pH 8.0). Enzyme (P450 or P411) and oxygen depletion mixture (20  $\mu$ L) were added to the vial before crimp-sealing. Portions of D-glucose (250 mM, 40  $\mu$ L), NADPH (100 mM, 40  $\mu$ L), and potassium phosphate buffer (0.1 M, pH = 8.0, appropriate volume so final reaction volume is 400  $\mu$ L), or multiples thereof, were combined in a larger crimp sealed vial and degassed by sparging with argon for at least 10 min. In the meantime, the headspace of the sealed 2 mL reaction vial with the enzyme solution was made anaerobic by flushing argon over the headspace (with no bubbling). The buffer/reductant/glucose solution (appropriate volume such that final reaction volume is 400  $\mu$ L) was added to the reaction vial via syringe under continuous argon purge of the vial headspace. Olefin (10  $\mu$ L, 300 mM in DMSO) then tosyl azide (10  $\mu$ L, 100 mM in DMSO) was added to the reaction vial using a glass syringe and the reaction was allowed to shake for 4 h at room temperature. Final reaction volume was 400  $\mu$ L; final concentrations of reagents were typically: 7.5 mM olefin, 2.5 mM tosyl azide, 10 mM NADPH, 25 mM D-glucose, and GOX oxygen depletion system. *Note:* Sodium dithionite (5 mM) was used as reductant in place of NADPH for reactions with hemin, myoglobin, cytochrome *c*, CYP119, and P450<sub>Rhf</sub>. After, reactions were quenched by adding acetonitrile (460  $\mu$ L) and the resulting mixture was transferred to a microcentrifuge tube and centrifuged at 14,000 rpm for 5 minutes. The solution (540  $\mu$ L) was transferred to an HPLC vial, charged with internal standard (60  $\mu$ L, 10 mM 1,3,5-trichlorobenzene in acetonitrile), and analyzed by HPLC. Concentrations of products in small-scale reactions were calculated using a calibration curve; TTN was calculated by dividing the concentration of

product by the concentration of protein catalyst.

### 2.5.7 Reaction screening in 96-well plate format

Site-saturation mutagenesis libraries were generated by employing the “22c-trick” method.<sup>37</sup> Libraries in *E. coli* were generated and cultured in 300  $\mu$ L LB<sub>amp</sub>. Pre-culture (50  $\mu$ L) was transferred to 1000  $\mu$ L of HB<sub>amp</sub> using a multichannel pipette. The cultures were incubated at 37 °C, 220 rpm, 80% humidity for 3 hours. The plates were cooled on ice for 15 minutes before expression was induced (0.5 mM IPTG, 1.0 mM 5-aminolevulinic acid, final concentrations). Expression was conducted at 20 °C, 120 rpm, 20 h. The cells were pelleted (3000 x g, 5 min) and re-suspended in 40  $\mu$ L/well GOX solution (14,000 U/ml catalase (Sigma 02071) and 1000 U/ml glucose oxidase (Sigma G7141)). The 96-well plate was transferred to an anaerobic chamber. In the anaerobic chamber, 300  $\mu$ L per well argon sparged reaction buffer (4 : 1, M9-N : 250 mM glucose in M9-N) was added followed by 4-methylstyrene (300 mM, 10  $\mu$ L/well) and tosyl azide (100 mM, 10  $\mu$ L/well). The plate was sealed with aluminum sealing tape, removed from the anaerobic chamber, and shaken at 400 rpm. After 16 hours, the seal was removed and 400  $\mu$ L of acetonitrile was added to each well. The contents of each well were mixed by pipetting up and down using a multichannel pipette. Then the plate was centrifuged (4,000 x g, 5 minutes) and 500  $\mu$ L of the supernatant was transferred to a shallow-well plate for analysis by HPLC.

### 2.5.8 Typical procedure for small-scale aziridination bioconversions under anaerobic conditions using P411 enzymes in *E. coli* cells.

*E. coli* BL21 *E. coli* cells carrying a plasmid encoding a P411 variant were inoculated from glycerol stock and grown overnight in 5 ml LB<sub>amp</sub> (37 °C, 250 rpm). The preculture was used to inoculate 45 mL of HB<sub>amp</sub> in a 125 mL Erlenmeyer flask; this culture was incubated at 37 °C, 220 rpm for 2 h and 30 min. After, the cultures were cooled on ice and induced with 0.5 mM IPTG and 1 mM 5-aminolevulinic acid (final concentration). Expression was conducted at room temperature, 130 rpm, 20 h. The cultures were then harvested and resuspended to OD<sub>600</sub> = 30 in M9-N. Aliquots of the cell suspension (4 mL) were used to determine the P450 or P411 expression level after lysis by sonication.



*E. coli* cells ( $OD_{600} = 30$ ) were made anaerobic by sparging with argon in a sealed 6 mL crimp vial for at least 30 minutes. To a 2 mL crimp vial was then added glucose (250 mM in M9-N, 40  $\mu$ L) and the GOX oxygen depletion solution described previously (20  $\mu$ L). The headspace of the sealed 2 mL reaction vial was made anaerobic by flushing argon over the solution. Resuspended cells (320  $\mu$ L), followed by olefin substrate (10  $\mu$ L, 300 mM in DMSO), then tosyl azide (10  $\mu$ L, 100 mM in DMSO) were added to 2 mL reaction vial via syringe under continuous flow of argon. Reactions were allowed to shake for 4 h at room temperature. Final reaction volume was 400  $\mu$ L; final concentrations of reagents were typically: 7.5 mM olefin, 2.5 mM tosyl azide, 25 mM D-glucose, and GOX oxygen depletion system. The no enzyme control experiment was conducted using *E. coli* BL21 cells containing empty pET22b(+) vector with the same reaction conditions as described above. The reactions were quenched by adding acetonitrile (460  $\mu$ L) and the resulting mixture was transferred to a microcentrifuge tube and centrifuged at 14,000 rpm for 5 minutes. The solution (540  $\mu$ L) was transferred to an HPLC vial, charged with internal standard (60  $\mu$ L, 10 mM 1,3,5-trichlorobenzene in acetonitrile), and analyzed by HPLC. Concentrations of products in small-scale reactions were calculated using a calibration curve; TTN was calculated by dividing the concentration of product by the concentration of protein catalyst.

Reactions for chiral HPLC analysis were performed on a 2 mL scale (final reaction volume) using the same concentration of reagents and a similar procedure as described above. Briefly, *E. coli* containing P411 enzymes were expressed and resuspended to an  $OD_{600} = 30$  in M9-N, and then degassed by sparging with argon in a sealed 6 mL crimp vial for at least 30 minutes. To a 6 mL crimp vial was then added glucose (250 mM in M9-N, 200  $\mu$ L) and the GOX oxygen depletion mixture described previously (100  $\mu$ L). The headspace of the sealed 2 mL reaction vial was made anaerobic by flushing argon over the solution. Resuspended cells (1600  $\mu$ L), followed by olefin substrate (50  $\mu$ L, 300 mM in DMSO), then tosyl azide (50  $\mu$ L, 100 mM in DMSO) were added to 6 mL reaction vial via syringe under continuous flow of argon. Reactions were allowed to shake for 4 h at room temperature. Reactions were quenched with 2 mL acetonitrile and extracted with ethyl

acetate, and concentrated under reduced pressure. The remaining material was resuspended in acetone (200  $\mu$ L) and purified by C18 semi-preparative HPLC. The purified product was concentrated, resuspended in acetonitrile, and analyzed by SFC for enantioselectivity.

### 2.5.9 Experimental methods for determination of initial rates

All initial rate experiments were conducted in an anaerobic chamber. Initial rate measurements were performed using 0.2 mol% purified enzymes in 400  $\mu$ L scale reactions. A sealed 6-mL vial charged with glucose (250 mM, 480  $\mu$ L), NADPH (100 mM, 480  $\mu$ L), and potassium phosphate buffer (0.1 M, pH = 8.0, 3240  $\mu$ L) was sparged for at least 30 minutes with argon. After the degassing was complete, the reaction solution, 2-mL vials charged with GOX oxygen depletion solution (20  $\mu$ L), and purified protein (250  $\mu$ M in potassium phosphate buffer), kept on ice, were brought into the anaerobic chamber. *Note:* the GOX oxygen depletion solution is a 20X stock solution containing 14,000 U/mL catalase, 1,000 U/mL glucose oxidase dissolved in 0.1 M KPi pH 8.0. The reaction solution (350  $\mu$ L) was added to each 2-mL vial and allowed to equilibrate in the anaerobic chamber for 30 minutes. Reaction vials were then placed on a shaker (400 rpm), charged with 10  $\mu$ L purified protein (250  $\mu$ M in 0.1 M KPi buffer, pH = 8.0) and 4-methyl styrene substrate (10  $\mu$ L, 300 mM in DMSO) followed by tosyl azide (10  $\mu$ L, 100 mM in DMSO). Reactions were set up in duplicate and products quantified at 1-2 minute intervals by quenching with acetonitrile (460  $\mu$ L). The resulting mixture was removed from the anaerobic chamber, transferred to a microcentrifuge tube and centrifuged at 14,000 rpm for 5 minutes. The solution (540  $\mu$ L) was transferred to an HPLC vial, charged with internal standard (60  $\mu$ L, 10 mM 1,3,5-trichlorobenzene in acetonitrile), and analyzed by HPLC. The rates of aziridination and azide reduction for different enzyme variants are presented in Table 2-3 & Figure A-3. The rate of azide reduction was determined in the presence of olefin **18** (7.5 mM).

### 2.6 References and notes

1. Degennaro, L.; Trinchera, P.; Luisi, R. Recent advances in the stereoselective synthesis of aziridines. *Chem. Rev.* **2014**, *114*, 7881–7929.

2. Sweeney, J. B. Aziridines: epoxides' ugly cousins? *Chem. Soc. Rev.* **2002**, *31*, 247–258.
3. Ohno, H. Synthesis and applications of vinylaziridines and ethynylaziridines. *Chem. Rev.* **2014**, *114*, 7784–7814.
4. Hsueh, N.; Clarkson, G. J.; Shipman, M. Generation and ring opening of aziridines in telescoped continuous flow processes. *Org. Lett.* **2015**, *17*, 3632–3635.
5. Ghorai, M. K.; Kumar, A.; Tiwari, D. P.  $\text{BF}_3 \cdot \text{OEt}_2$ -Mediated highly regioselective  $\text{S}_{\text{N}}2$ -type ring-opening of *N*-activated aziridines and *N*-activated azetidines by tetraalkylammonium halides. *J. Org. Chem.* **2010**, *75*, 137–151.
6. Das, B.; Reddy, V. S.; Krishnaiah, M.; Rao, Y. K. Highly regio- and stereoselective ring-opening of epoxides and aziridines with sodium azide using ammonium-12-molybdophosphate. *J. Mol. Catal. A: Chem.* **2007**, *270*, 89–92.
7. Bera, M.; Roy, S. Silver(I)-diene complexes as versatile catalysts for the *C*-arylation of *N*-tosylaziridines: Mechanistic insight from in situ diagnostics. *J. Org. Chem.* **2010**, *75*, 4402–4412.
8. Huang, C.-Y.; Doyle, A. G. Nickel-catalyzed Negishi alkylations of styrenyl aziridines. *J. Am. Chem. Soc.* **2012**, *134*, 9541–9544.
9. Chandrasekhar, S.; Ahmed, M. Reductive opening of aziridines with polymethylhydrosiloxane. *Tetrahedron Lett.* **1999**, *40*, 9325–9327.
10. Bhandra, S. et al. Alumina-supported Cu(II), a versatile and recyclable catalyst for regioselective ring opening of aziridines and epoxides and subsequent cyclization to functionalized 1,4-benzoxazines and 1,4-benzodioxanes. *J. Org. Chem.* **2010**, *75*, 8533–8541.
11. Chai, Z.; Zhu, Y.-M.; Yang, P.-J.; Wang, S.; Wang, S.; Liu, Z.; Yang, G. Copper(I)-catalyzed kinetic resolution of *N*-sulfonylaziridines with indoles: Efficient construction of pyrroloindolines. *J. Am. Chem. Soc.* **2015**, *137*, 10088–10091.
12. Gao, G.-Y.; Jones, J. E.; Vyas, R.; Harden, J. D.; Zhang, X. P. Cobalt-catalyzed aziridination with diphenylphosphoryl azide (DPPA): direct synthesis of *N*-phosphorus-substituted aziridines from alkenes. *J. Org. Chem.* **2006**, *71*, 6655–6658.
13. Ruppel, J. V.; Jones, J. E.; Huff, C. A.; Kamble, R. M.; Chen, Y.; Zhang, X. P. A highly effective cobalt catalyst for olefin aziridination with azides: hydrogen bonding guided catalyst design. *Org. Lett.* **2008**, *10*, 1995–1998.

- 
14. Jin, L. M.; Xu, X.; Lu, H.; Cui, X.; Wojtas, L.; Zhang, X. P. Effective synthesis of chiral N-fluoroaryl aziridines through enantioselective aziridination of alkenes with fluoroaryl azides. *Angew. Chem. Int. Ed.* **2013**, *52*, 5309–5313.
  15. Lui, Y.; Che, C. M. [FeIII(F20-tpp)Cl] is an effective catalyst for nitrene transfer reactions and amination of saturated hydrocarbons with sulfonyl and aryl azides as nitrogen source under thermal and microwave-assisted conditions. *Chem. Eur. J.* **2010**, *16*, 10494–10501.
  16. Mansuy, D.; Mahy, J. P.; Dureault, A.; Bedi, G.; Battioni, P. Iron- and manganese-porphyrin catalysed aziridination of alkenes by tosyl- and acyl-iminoiodobenzene. *J. Chem. Soc., Chem. Commun.* **1984**, *17*, 1161–1163.
  17. Vyas, R.; Goa, G.-Y.; Harden, J. D.; Zhang, X. P. Iron(III) porphyrin catalyzed aziridination of alkenes with bromamine-T as nitrene source. *Org. Lett.* **2004**, *6*, 1907–1910.
  18. Ogasawara, Y.; Liu, H.-w. Biosynthetic studies of aziridine formation in azicemicins. *J. Am. Chem. Soc.* **2009**, *131*, 18066–18068.
  19. Tsutsumi, H.; Katsuyama, Y.; Izumikawa, M.; Takagi, M.; Fujie, M.; Satoh, N.; Shinya, K.; Ohnishi, Y. Unprecedented cyclization catalyzed by a cytochrome P450 in benzastatin biosynthesis. *J. Am. Chem. Soc.* **2018**, *140*, 6631–6639.
  20. Farwell, C. C. et al. Enantioselective imidation of sulfides via enzyme-catalyzed intermolecular nitrogen-atom transfer. *J. Am. Chem. Soc.* **2014**, *136*, 8766–8771.
  21. McIntosh, J. A.; et al. Enantioselective intramolecular C–H amination catalyzed by engineered cytochrome P450 enzymes in vitro and in vivo. *Angew. Chem. Int. Ed.* **2013**, *52*, 9309–9312.
  22. Coelho, P. S.; et al. A serine-substituted P450 catalyzes highly efficient carbene transfer to olefins in vivo. *Nat. Chem. Biol.* **2013**, *9*, 485–487.
  23. Yosca, T. H.; et al. Iron(IV)hydroxide pKa and the role of thiolate ligation in C–H bond activation by cytochrome P450. *Science* **2013**, *342*, 825–829.
  24. Perera, R.; et al. Molecular basis for the inability of an oxygen atom donor ligand to replace the natural sulfur donor heme axial ligand in cytochrome P450 catalysis. Spectroscopic characterization of the Cys436Ser CYP2B4 mutant. *Arch. Biochem. Biophys.* **2011**, *507*, 119–125.
  25. Hyster, T. K.; et al. Enzyme-controlled nitrogen-atom transfer enables regiodivergent C–H amination. *J. Am. Chem. Soc.* **2014**, *136*, 15505–15508.

- 
26. Ando, T.; et al. Iodine-catalyzed aziridination of alkenes using chloramine-T as a nitrogen source. *Tetrahedron* **1998**, *54*, 13485–13494.
27. Heel, T.; et al. Non-natural olefin cyclopropanation catalyzed by diverse cytochrome P450s and other hemoproteins. *ChemBioChem* **2014**, *15*, 2556–2562.
28. Wang, Z. J.; et al. Improved cyclopropanation activity of histidine-ligated cytochrome P450 enables the enantioselective formal synthesis of levomilnacipran. *Angew. Chem. Int. Ed.* **2014**, *53*, 6810–6813.
29. The following observation was made in reference 20. When a solution of tosyl azide, degassed with argon, was added to a ferrous P411<sub>BM3</sub> variant in purified form, an immediate shift back to the ferric state was observed with concomitant production of *p*-toluenesulfonamide. This suggests that following *p*-toluenesulfonamide formation, ferrous P411<sub>BM3</sub> returns to the ferric state.
30. Tao, J. Asymmetric nitrene transfer reactions with azides via Co(II)-based metalloradical catalysis (MRC). *University of South Florida Doctoral Dissertation*, **2013**. Retrieved from University of South Florida, available here: <https://scholarcommons.usf.edu/etd/4590/>.
31. Zhu, D. et al. Inverting the enantioselectivity of a carbonyl reductase via substrate-enzyme docking-guided point mutation. *Org. Lett.* **2008**, *10*, 525–528.
32. Guo, C. et al. Exploring the enantioselective mechanism of halohydrin dehalogenase from *Agrobacterium radiobacter* AD1 by iterative saturation mutagenesis. *Appl. Environ. Microbiol.* **2015**, *81*, 2919–2926.
33. Bornscheuer, U. T.; Kazlauskas, R. J. Catalytic promiscuity in biocatalysis: using old enzymes to form new bonds and follow new pathways. *Angew. Chem. Int. Ed.* **2004**, *43*, 6032–6040.
34. Renata, H.; Wang, Z. J.; Arnold, F. H. Expanding the enzyme universe: accessing non-natural reactions by mechanism guided directed evolution. *Angew. Chem. Int. Ed.* **2015**, *54*, 3351–3367.
35. Srinivas, B.; Kumar, V.P.; Sridhar, R.; Surendra, K.; Nageswar, Y.V.D; Rao, K.R. *J. Mol. Catal. A: Chem.* **2007**, *261*, 1–5.
36. Ando, T.; Kano, D.; Minakata, S.; Ryu, I; Komatsu, M. *Tetrahedron* **1998**, *54*, 13485–13494
37. Kille, S., Acevedo-Rocha, C. G., Parra, L. P., Zhang, Z.-G., Opperman, D. J., Reetz, M. T., Acevedo J. P. Reducing codon redundancy and screening effort of

- 
- combinatorial protein libraries created by saturation mutagenesis. *ACS Synth. Biol.* **2013**, *2*, 83–92.
38. Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchinson III, C. A., Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **2009**, *6*, 343–345.
39. Omura T.; Sato, R. The Carbon Monoxide-binding Pigment of liver microsomes: I. Evidence for its hemoprotein nature. *J. Biol. Chem.* **1964**, *239*, 2370–2378.
40. Vatsis, K. P., Peng, H.-M., Coon, M. J. Replacement of active-site cysteine-436 by serine converts cytochrome P450 2B4 into an NADPH oxidase with negligible monooxygenase activity. *J. Inorg. Biochem.* **2002**, *91*, 542–553.
41. Fuhrhop, J.-H., Smith, K. M. Laboratory methods in porphyrin and metalloporphyrin research. (Elsevier, 1975).

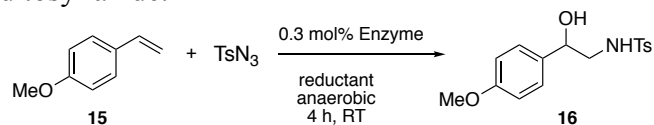
## SUPPLEMENTARY INFORMATION FOR CHAPTER 2

Material for this chapter appears in Farwell, C. C.<sup>†</sup>; **Zhang, R. K.**<sup>†</sup>; McIntosh, J. A.; Hyster, T. K.; Arnold F. H. “Enantioselective enzyme-catalyzed aziridination enabled by active-site evolution of a cytochrome P450,” *ACS Central Science* **2015**, *1*, 89–93. **DOI:** 10.1021/acscentsci.5b00056. (<sup>†</sup>Denotes equal contribution) This work was performed in collaboration with all authors. Reprinted with permission from American Chemical Society.

**A.1 Supporting Tables A-1 through A-5****Table A-1.** Mutations present in P450<sub>BM3</sub> variants used in Chapter 2.

Enzyme	Mutations relative to wild type P450 <sub>BM3</sub>
<b>P450<sub>BM3</sub></b>	none
<b>P450<sub>BM3</sub> T268A</b>	T268A
<b>P411<sub>BM3</sub></b>	C400S
<b>P411<sub>BM3</sub> T268A</b>	T268A, C400S
<b>P450<sub>BM3</sub>-CIS T438S</b>	V78A, F87V, P142S, T175I, A184V, S226R, H236Q, E252G, T268A, A290V, L353V, I366V, T438S, E442K
<b>P411<sub>BM3</sub>-CIS T438S</b>	P450 <sub>BM3</sub> -CIS C400S, T438S
<b>P-I263F</b>	P411 <sub>BM3</sub> -CIS T438S I263F
<b>P-I263F A328V</b>	P411 <sub>BM3</sub> -CIS T438S I263F A328V
<b>P-I263F A328V L437V</b>	P411 <sub>BM3</sub> -CIS T438S I263F A328V L437V
<b>P411<sub>BM3</sub>-CIS A268T T438S</b>	P411 <sub>BM3</sub> -CIS A268T, C400S, T438S
<b>P411-H2A10</b>	P411 <sub>BM3</sub> -CIS L75A, L181A
<b>P411-H2A10 I263F</b>	P411 <sub>BM3</sub> -CIS L75A, L181A, I263F
<b>P411<sub>BM3</sub> H2-5-F10</b>	P411 <sub>BM3</sub> -CIS L75A, I263A, L437A
<b>P411<sub>BM3</sub> H2-4-D4</b>	P411 <sub>BM3</sub> -CIS L75A, M177A, L181A, L437A

**Table A-2.** Heme and other heme proteins tested for aziridination activity with 4-methoxystyrene and tosyl azide.<sup>a</sup>

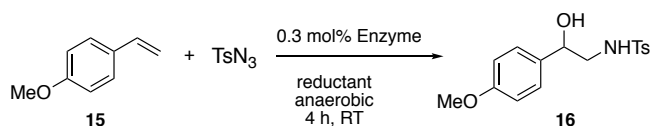


Entry	Catalyst <sup>b</sup>	TTN 2
1	Hemin	< 1
2	Hemin + BSA	< 1
3	Myoglobin (horse heart)	< 1
4	Cytochrome c (bovine heart)	< 1
5	CYP119 C317S	7
6	CYP119 T213A C317H	< 1
7	P450 <sub>Rhf</sub>	< 1
8	P450 <sub>Rhf</sub> T275A	< 1

<sup>a</sup> Purified enzyme reactions were conducted as described in the experimental methods section (Chapter 2.5), using 7.5 mM olefin and 2.5 mM TsN<sub>3</sub>.<sup>1</sup> Sodium dithionite (5 mM) was used as reductant. <sup>b</sup> Myoglobin and cytochrome *c* were purchased as lyophilized powder from Sigma Aldrich. P450<sub>Rhf</sub> mutants were expressed and purified as described in the methods section; P450<sub>CYP119</sub> was expressed and purified as described previously.<sup>2</sup>

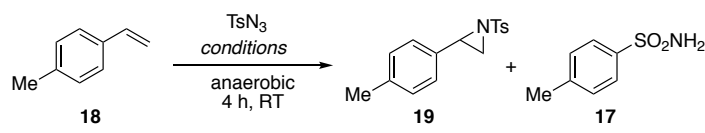


**Table A-3.** Panel of P450<sub>BM3</sub> and P411 purified enzymes tested for aziridination activity with 4-methoxystyrene and tosyl azide.<sup>a, b</sup>



Entry	Enzyme	TTN 16
1	P411 <sub>BM3</sub> -CIS T438S	15
2	P450 <sub>BM3</sub> -CIS T438S	< 1
3	P450 <sub>BM3</sub> -CIS T438S C400H	3
4	P450 <sub>BM3</sub> -CIS T438S C400D	4
5	P450 <sub>BM3</sub> -CIS T438S C400M	4
6	P411 <sub>BM3</sub> -CIS A268T T438S	< 1
7	P411 <sub>BM3</sub> -H2-5-F10	8
8	P411 <sub>BM3</sub> -H2-A-10	4
9	P411 <sub>BM3</sub> -H2-4-D4	5
10	P450 <sub>BM3</sub>	< 1
11	P411 <sub>BM3</sub>	3
12	P450 <sub>BM3</sub> -T268A	2
13	P411 <sub>BM3</sub> -T268A	4
14	P411 <sub>BM3</sub> -CIS T438S I263F	150
15	P411 <sub>BM3</sub> -CIS T438S I263F V87F	19
16	P411 <sub>BM3</sub> -CIS T438S I263F A268T	< 1

<sup>a</sup>Purified enzyme reactions were conducted as described in the experimental methods section (Chapter 2.5). <sup>b</sup>Reactions were conducted by Dr. Chris Farwell.

**Table A-4.** Control experiments using purified P-I263F proteins and NADPH as reductant.<sup>a</sup>


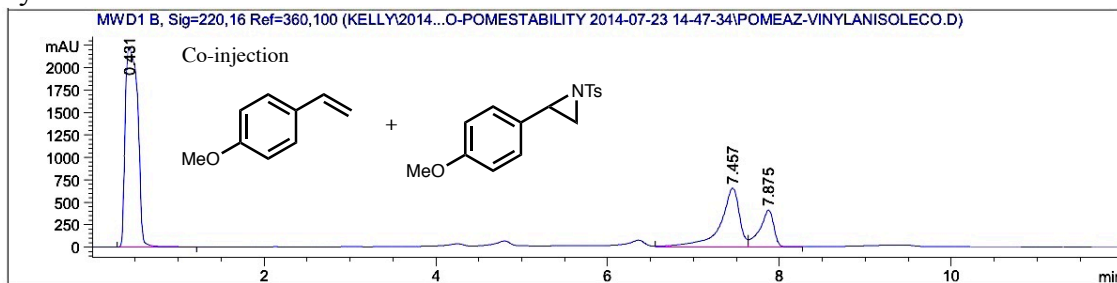
Entry	Conditions	mM <b>19</b>	mM <b>17</b>
1	substrates only	N.D.	0.02
2	substrates and NADPH	N.D.	0.14
3	P-I263F and substrates	N.D.	0.03
4	P-I263F, substrates, and NADPH	0.92	1.40
5	P-I263F heme and substrates	N.D.	0.02
6	P-I263F heme, substrates, and NADPH	N.D.	0.27

<sup>a</sup> Concentrations are 7.5 mM olefin, 2.5 mM TsN<sub>3</sub>, 10 mM NADPH (when applicable), 5.0 μM P-I263F proteins (when applicable) in M9-N buffer. P-I263F refers to full-length P-I263F enzyme. P-I263F heme refers to the heme domain only of the protein (without the reductase domain). The mixtures were allowed to shake for 4 h at room temperature under anaerobic conditions to mimic conditions used for *in vitro* (purified protein) experiments. N.D., not detected.

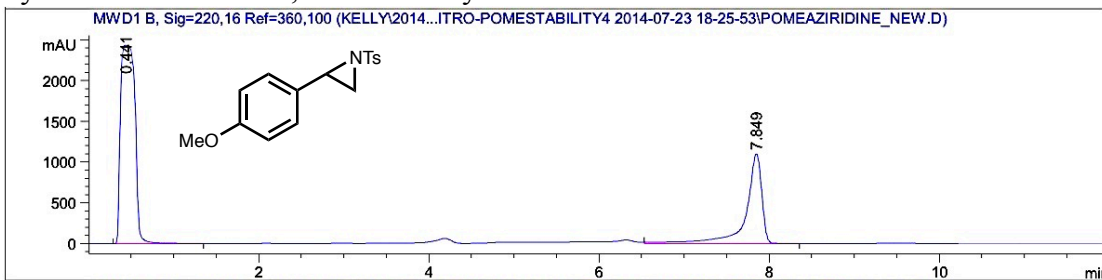
## A.2 Supporting Figures A-1 through A-3 and Related Experimental Details

(a) HPLC chromatogram of controls, monitored at 220 nm.

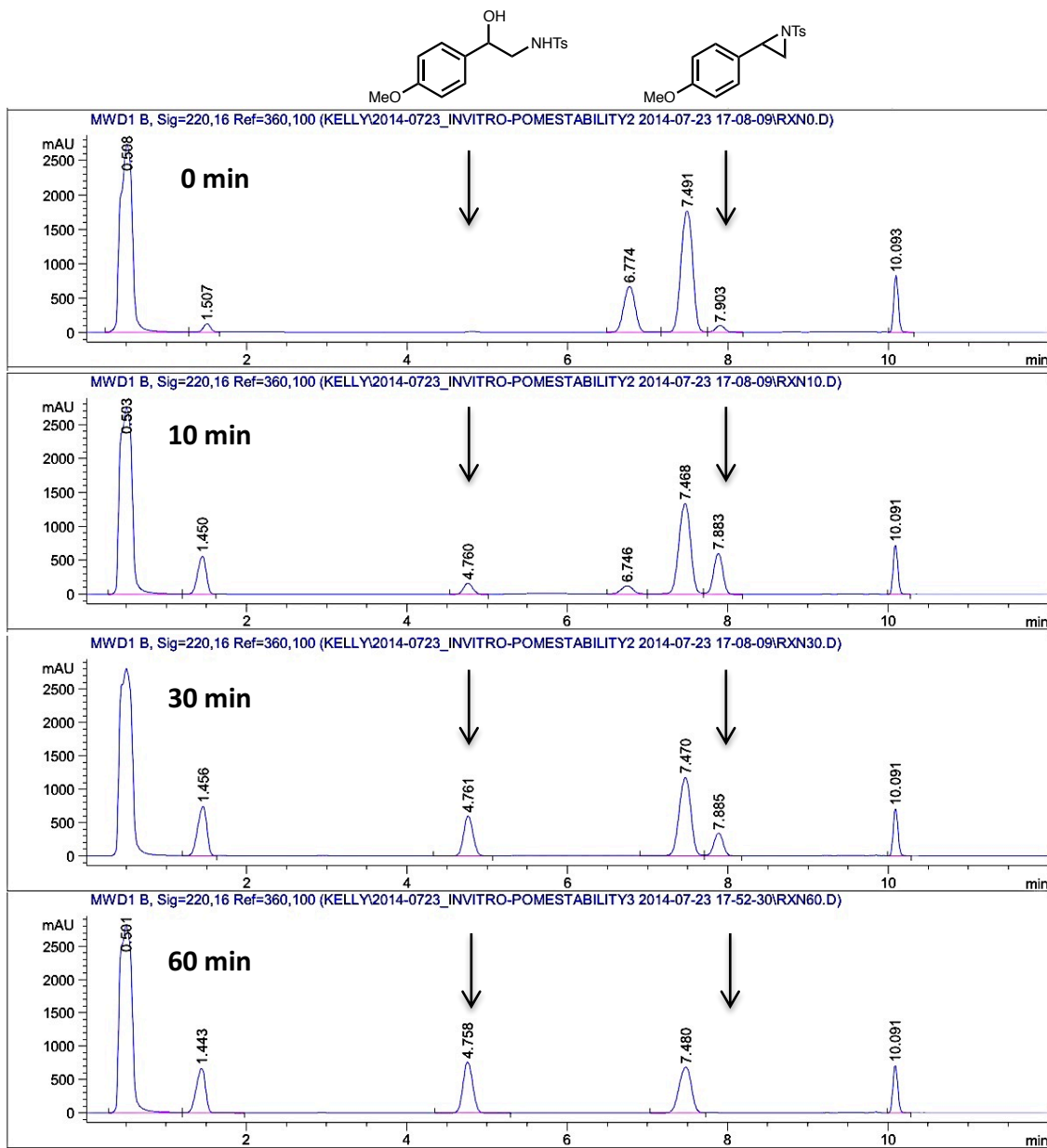
Co-injection of 4-methoxystyrene (Sigma Aldrich) and synthetic standard **16'**, confirmed by NMR:



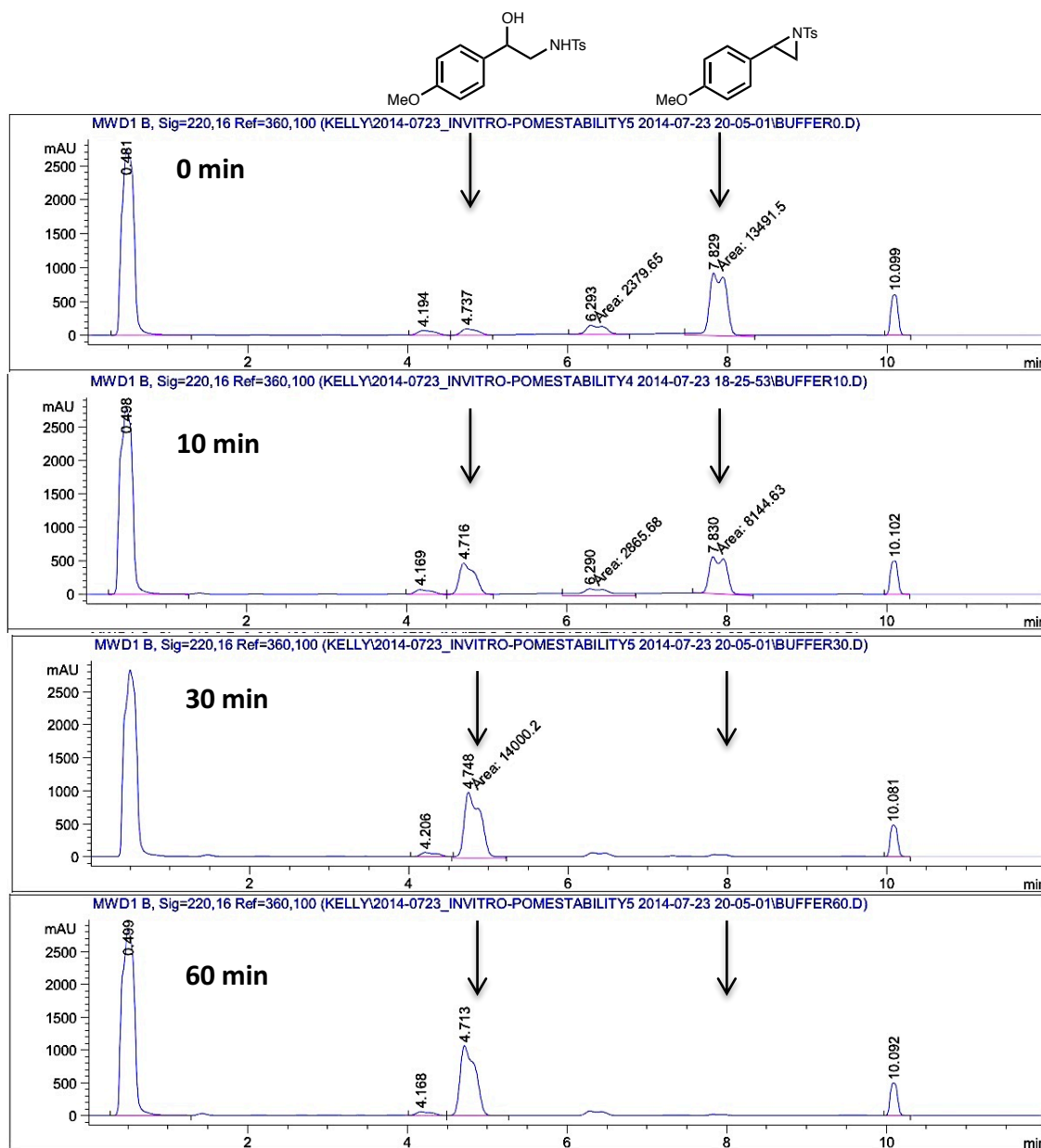
Synthetic standard **16'**, confirmed by NMR:



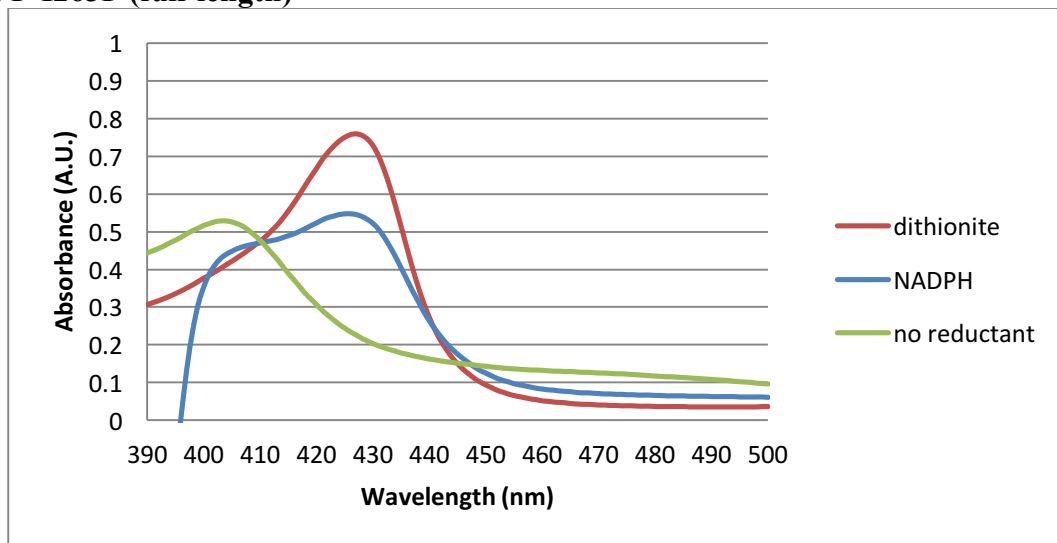
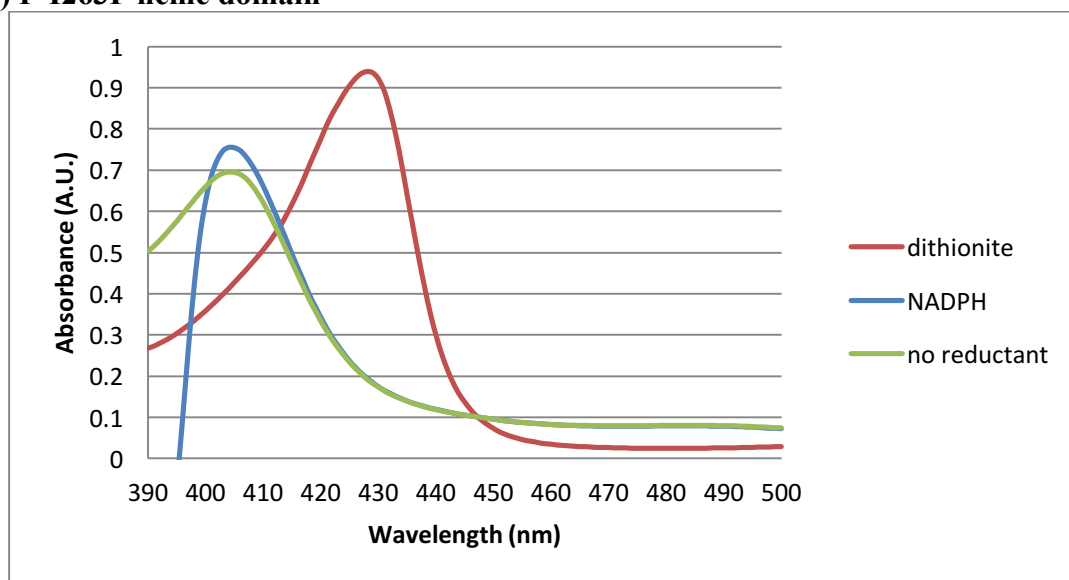
(b) HPLC chromatograms of P411-enzymatic reaction with 4-methoxystyrene **15** and tosyl azide as substrates analyzed at different time points, monitored at 220 nm. Putative aziridine **16'** and amido-alcohol **16** are marked with arrows.



(c) HPLC chromatograms of synthetic standard **16'** in reaction conditions *without* P411 catalyst at several time points, monitored at 220 nm. Putative aziridine **16'** and amido-alcohol **16** are marked with arrows.



**Figure A-1.** Demonstration of enzymatic synthesis and degradation of aziridine **16'** under reaction conditions.

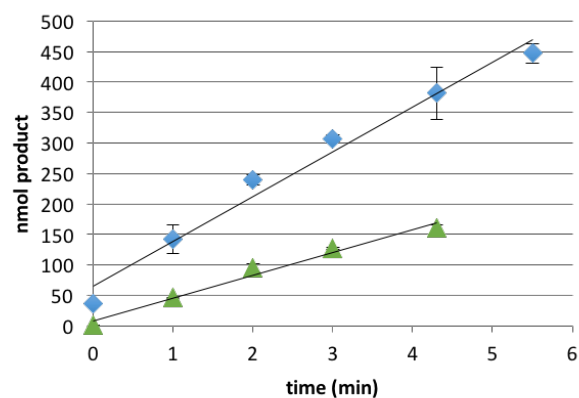
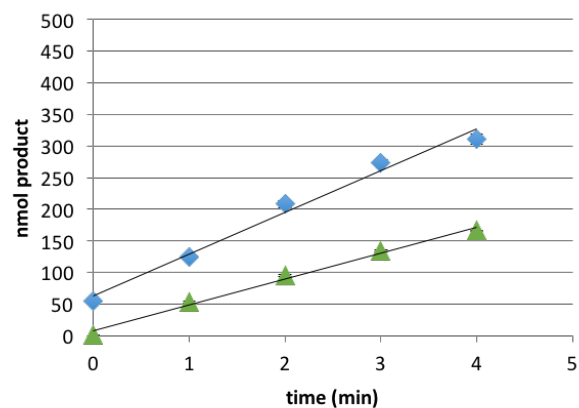
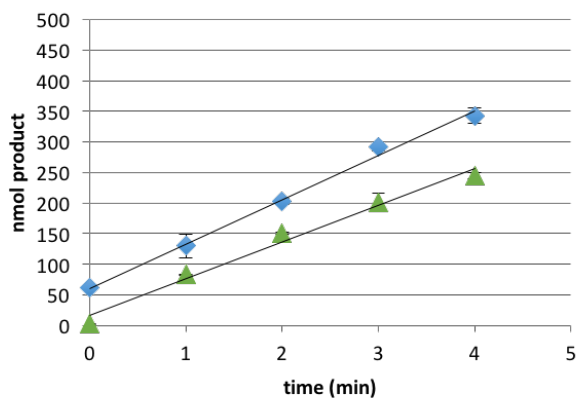
**(a) P-I263F (full-length)****(b) P-I263F heme domain**

**Figure A-2.** UV-vis absorbance spectra of P-I263F (full-length) (a) and P-I263F heme domain only proteins (b) after addition of NADPH. Representative UV-vis absorbance spectra are shown for purified protein in the presence of no reductant ( $\text{Fe}^{\text{III}}$ , green), NADPH (blue), and dithionite ( $\text{Fe}^{\text{II}}$ , red). In the case of full-length P-I263F, the  $\text{Fe}^{\text{II}}$  (426 nm) Soret band is observed when either NADPH or dithionite is used as reductant. The shoulder at 404 nm observed in the NADPH spectrum is due to incomplete reduction of full-length P-I263F under the experimental conditions. For P-I263F heme domain, only the  $\text{Fe}^{\text{III}}$  (404 nm) Soret band is observed when NADPH is used as the reductant.

**Experimental methods for UV-visible absorbance spectroscopy results presented in Figure A-2.**

UV-Visible absorbance spectroscopy was performed under anaerobic conditions. A sealed 6-mL vial charged with potassium phosphate buffer (0.1 M, pH = 8.0, 4 mL) and a sealed 2-mL vial charged with NADPH (100 mM in 0.1 M KPi pH = 8.0 buffer, 1 mL) were sparged for at least 20 minutes with argon. In parallel, purified full-length P-I263F (200  $\mu$ M, 25  $\mu$ L) or PI263F heme domain only (250  $\mu$ M, 20  $\mu$ L) was added to a semi-micro cuvette. The cuvette was sealed with a cap equipped with rubber septa and the headspace was purged with argon for at least 10 minutes. After degassing was complete, potassium phosphate buffer (880  $\mu$ L), followed by NADPH (100 mM, 100  $\mu$ L), were added to the anaerobic cuvette containing protein via syringe under a continuous stream of argon. UV-vis spectra of the protein sample was recorded until a stable ferrous state was reached, or for 20 minutes if no ferrous state was observed.

The negative control (no reductant) and positive control (dithionite-reduced protein) experiments were performed in a similar manner except degassed potassium phosphate buffer (100  $\mu$ L, negative control) or degassed dithionite solution (100 mM in 0.1 M Kpi pH = 8.0, 100  $\mu$ L, positive control) was added to the protein sample instead of NADPH solution.

**(a) P-I263F****(b) P-I263F A328V****(c) P-I263F A328V L437V**

**Figure A-3.** Data used to determine initial rate results presented in Table 2-3. Blue diamonds represent sulfonamide **17** and green triangles represent aziridine **19** for all plots.

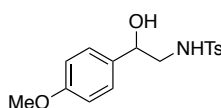


### A.3 Synthesis and Characterization of Substrates and Reference Compounds

All olefin substrates presented in Table 2-3 were obtained from commercial sources (Sigma Aldrich, Alfa Aesar) and used as received. Racemic reference compounds corresponding to enzymatic products were prepared according to the following procedures. Reference compounds are characterized below.

#### *N*-(2-hydroxy-2-(4-methoxyphenyl)ethyl)-4-methylbenzenesulfonamide (16).

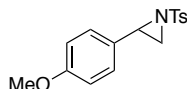
This compound was prepared following the method of B. Srinivas et al.<sup>3</sup>



<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.72 (d, 2H, *J* = 8.1 Hz), 7.29 (d, 2H, *J* = 8.3 Hz), 7.19 (d, 2H, *J* = 8.6 Hz), 6.84 (d, 2H, *J* = 8.6 Hz), 5.06 (dd, 1H, *J* = 8.1, 4.6 Hz), 4.73 (dd, 1H, *J* = 8.7, 3.7 Hz), 3.78 (s, 3H), 3.20 (ddd, 1H, *J* = 13.3, 8.1, 3.7 Hz), 3.01 (ddd, 1H, *J* = 13.2, 8.6, 4.6 Hz), 2.42 (s, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 159.66, 143.69, 136.86, 133.00, 129.90, 127.26, 127.21, 114.16, 72.50, 55.44, 50.30, 21.66. HRMS (FAB<sup>+</sup>): calculated for C<sub>16</sub>H<sub>18</sub>NO<sub>4</sub>S ([M+H]<sup>+</sup>): 320.0956; found: 320.0950

#### *N*-(*p*-Tolylsulfonyl)-2-(*p*-methoxyphenyl)aziridine (16').

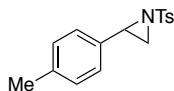
This compound was prepared following the method of T. Ando et al.<sup>4</sup>; spectral data are in agreement with literature values.<sup>5</sup>



<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.87 (d, 2H, *J* = 8.3 Hz), 7.34 (d, 2H, *J* = 8.5 Hz), 7.14 (d, *J* = 8.7 Hz, 2H), 6.83 (d, *J* = 8.7, 2H), 3.78 (s, 3H), 3.75 (dd, 1H, *J* = 7.2, 4.5 Hz), 2.97 (d, 1H, *J* = 7.2 Hz), 2.44 (s, 3H), 2.39 (d, 1H, *J* = 4.5 Hz).

#### *N*-(*p*-Tolylsulfonyl)-2-(*p*-methylphenyl)aziridine (19).

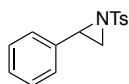
This compound was prepared following the method of T. Ando et al.<sup>4</sup>; spectral data are in agreement with literature values.<sup>5</sup>



<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.86 (d, 2H, *J* = 8.3 Hz), 7.32 (d, 2H, *J* = 8.3 Hz), 7.10 (s, 4H), 3.74 (dd, 1H, *J* = 7.2, 4.5 Hz), 2.97 (d, 1H, *J* = 7.2 Hz), 2.43 (s, 3H), 2.38 (d, 1H, *J* = 4.5 Hz), 2.31 (s, 3H).

#### *N*-(*p*-Tolylsulfonyl)-2-phenylaziridine (21).

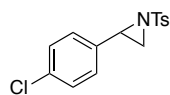
This compound was prepared following the method of T. Ando et al.<sup>4</sup>; spectral data are in agreement with literature values.<sup>6</sup>



<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.87 (d, 2H, *J* = 8.3 Hz), 7.19-7.36 (m, 7H), 3.77 (dd, 1H, *J* = 7.2, 4.5 Hz), 2.98 (d, 1H, *J* = 7.2 Hz), 2.43 (s, 3H), 2.39 (d, 1H, *J* = 4.5 Hz).

***N*-(*p*-Tolylsulfonyl)-2-(*p*-chlorophenyl)aziridine (22).**

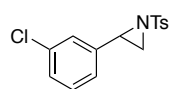
This compound was prepared following the method of T. Ando et al.<sup>4</sup>; spectral data are in agreement with literature values.<sup>6</sup>



**<sup>1</sup>H NMR** (300 MHz, CDCl<sub>3</sub>) δ 7.86 (d, 2H, *J* = 8.3 Hz), 7.34 (d, 2H, *J* = 7.9 Hz), 7.26 (d, 2H, *J* = 8.5 Hz), 7.15 (d, 2H, *J* = 8.5 Hz), 3.73 (dd, 1H, *J* = 7.2, 4.4 Hz), 2.98 (d, 1H, *J* = 7.2 Hz), 2.44 (s, 3H), 2.34 (d, 1H, *J* = 4.4 Hz).

***N*-(*p*-Tolylsulfonyl)-2-(*m*-chlorophenyl)aziridine (23).**

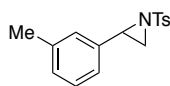
This compound was prepared following the method of T. Ando et al.<sup>4</sup>; spectral data are in agreement with literature values.<sup>7</sup>



**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.87 (d, 2H, *J* = 8.3 Hz), 7.35 (d, 2H, *J* = 7.7 Hz), 7.19–7.26 (m, 3H), 7.12 (dt, 1H, *J* = 6.8, 1.8 Hz), 3.73 (dd, 1H, *J* = 7.2, 4.3 Hz), 2.97 (d, 1H, *J* = 7.2 Hz), 2.44 (s, 3H), 2.35 (d, 1H, *J* = 4.4 Hz).

***N*-(*p*-Tolylsulfonyl)-2-(*m*-methylphenyl)aziridine (24).**

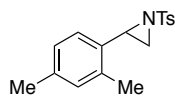
This compound was prepared following the method of T. Ando et al.<sup>4</sup>; spectral data are in agreement with literature values.<sup>8</sup>



**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.87 (d, 2H, *J* = 8.3 Hz), 7.33 (d, 2H, *J* = 8.6 Hz), 7.01–7.20 (m, 4H), 3.74 (dd, 1H, *J* = 7.2, 4.5 Hz), 2.96 (d, 1H, *J* = 7.2 Hz), 2.43 (s, 3H), 2.38 (d, 1H, *J* = 4.5 Hz), 2.30 (s, 3H).

***N*-(*p*-Tolylsulfonyl)-2-(2,4-dimethylphenyl)aziridine (25).**

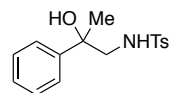
This compound was prepared following the method of T. Ando et al.<sup>4</sup>



**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.90 (d, 2H, *J* = 8.4 Hz), 7.34 (d, 2H, *J* = 8.5 Hz), 6.91–7.00 (m, 3H), 3.84 (dd, 1H, *J* = 7.2, 4.6 Hz), 2.97 (d, 1H, *J* = 7.2 Hz), 2.44 (s, 3H), 2.35 (s, 3H), 2.32 (d, 1H, *J* = 4.6 Hz), 2.28 (s, 3H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 144.72, 137.95, 136.72, 135.15, 130.89, 130.32, 129.84, 128.11, 126.82, 125.98, 39.61, 35.07, 21.75, 21.11, 19.08. **HRMS** (FAB<sup>+</sup>): calculated for C<sub>17</sub>H<sub>20</sub>NO<sub>2</sub>S ([M+H]<sup>+</sup>): 302.1215; found: 302.1210.

***N*-(2-hydroxy-2-phenylpropyl)-4-methylbenzenesulfonamide (26).**

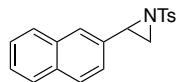
This compound was prepared following the method of B. Srinivas et al.<sup>3</sup>



**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.67 (d, 2H, *J* = 8.3 Hz), 7.24–7.38 (m, 7H), 4.59 (s, 1H), 3.22 (dd, 1H, *J* = 12.8, 8.5 Hz), 3.12 (dd, 1H, *J* = 12.8, 4.8 Hz), 2.42 (s, 3H), 1.56 (s, 3H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 144.87, 143.73, 136.73, 129.93, 128.75, 127.60, 127.19, 124.93, 73.81, 53.99, 27.62, 21.68. **HRMS** (FAB<sup>+</sup>): calculated for C<sub>16</sub>H<sub>20</sub>NO<sub>3</sub>S ([M+H]<sup>+</sup>): 306.1164; found: 306.1160.

***N*-(*p*-Tolylsulfonyl)-2-(2-naphthyl)aziridine (27).**

This compound was prepared following the method of T. Ando et al.<sup>4</sup>; spectral data are in agreement with literature values.<sup>6</sup>



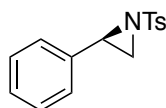
**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$  7.90 (d, 2H,  $J$  = 8.3 Hz), 7.75 – 7.81 (m, 3H), 7.73 (s, 1H), 7.45– 7.49 (m, 2H), 7.33 (d, 2H,  $J$  = 8.3 Hz), 7.25 – 7.30 (m, 1H), 3.93 (dd, 1H,  $J$  = 7.2, 4.4 Hz), 3.07 (d, 1H,  $J$  = 7.2 Hz), 2.50 (d, 1H,  $J$  = 4.5 Hz), 2.42 (s, 3H).

**A.4 Calibration Curves**

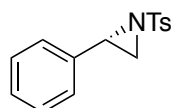
Calibration curves with an internal standard were created for quantitative HPLC analysis of reaction products; these curves plot the response factor (the ratio of product area to internal standard area, y-axis) against the concentration of product (x-axis). Internal standard was 1,3,5-trichlorobenzene (1.0 mM) and reactions were monitored at 220 nm. Calibration curves can be found in the Supplementary Information of the published paper. The identity of the products was additionally confirmed by LC-MS co-injections of reaction mixtures with chemically synthesized reference compounds; these traces can be found in the Supporting Information of the published paper.

### A.5 Assignment of Absolute Stereochemistry

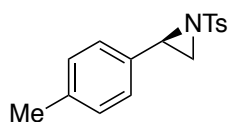
Absolute stereochemistry of enzymatically produced aziridine **21** was assigned by chiral HPLC analysis and optical rotation. In particular, absolute stereochemistry of **21** was previously assigned by chiral HPLC using Chiracel OJ column (isopropanol/ n-hexane mobile phase), with (*S*)-**21** the earlier eluting enantiomer.<sup>9</sup> Analytically enantiopure **21** produced by P-I263F A328V L437V (using whole cell catalyst) was subjected to the same chiral HPLC conditions and observed to be the earlier eluting enantiomer (Figure A-4), leading to an assignment of (*S*)-**21**. Further support for this assignment came from measuring optical rotation. The optical rotation values for enantiomers of **21** have been previously reported: (*R*)-**21**  $[\alpha]_D^{24} = -80.25^\circ$  (*c* 0.8, CHCl<sub>3</sub>)<sup>10</sup> and (*S*)-**21**  $[\alpha]_D^{20} = +26.7^\circ$  (*c* 0.7, CHCl<sub>3</sub>).<sup>11</sup> Optical rotation measurement of analytically enantiopure **21** produced by P-I263F A328V L437V gave  $[\alpha]_D^{25} = +80.2^\circ$  (*c* 1.2, CHCl<sub>3</sub>), revealing it to be (*S*)-**21**. Similarly, the optical rotation of P-I263F-A328V-L437V produced **19** (analytically enantiopure) was measured to be  $[\alpha]_D^{25} = +106.1^\circ$  (*c* 0.45, CHCl<sub>3</sub>). By analogy, the configuration of enzymatically preferred (+)-**19** is assigned as (*S*)-**19**.



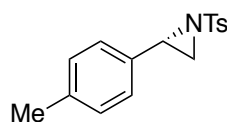
(-)-**21**



(+)-**21**

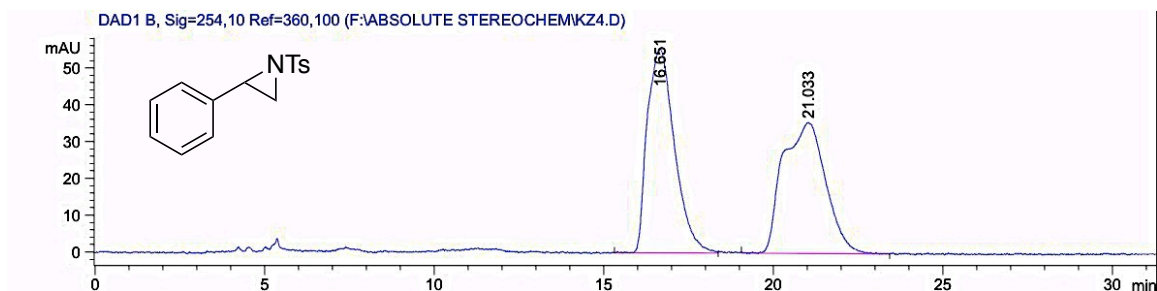


(-)-**19**

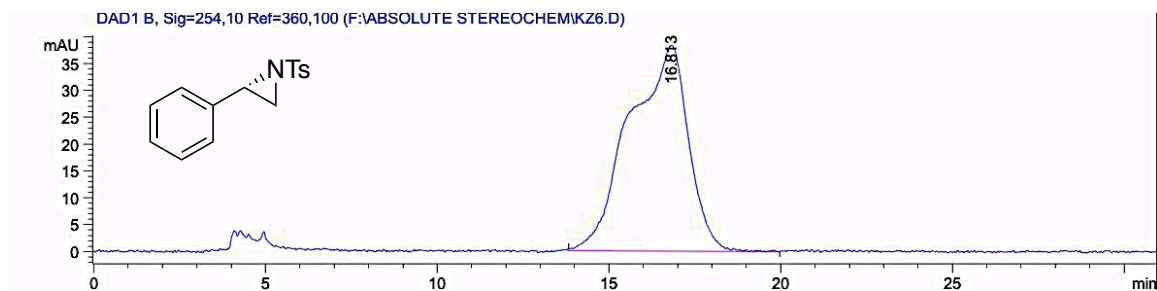


(+)-**19**

**Racemic synthetic aziridine 21,  $t_R = 16.7$  min and 21.0 min.**



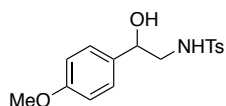
**P-I263F A328V L437V produced aziridine 21,  $t_R = 16.8$  min.**



**Figure A-4.** Assignment of absolute stereochemistry of enzymatically produced aziridine **21** by chiral HPLC (Chiracel OJ column, 30% isopropanol : 70% n-hexane, 210 nm).

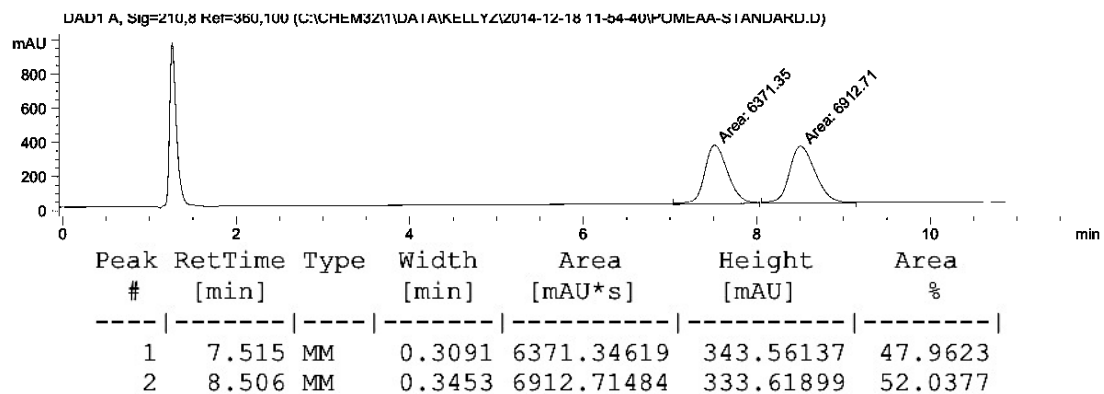
#### A.6 Determination of Enantioselectivity

Enantioselectivities of enzymatic products were determined by chiral SFC analysis. Representative traces, and their conditions, are shown below.

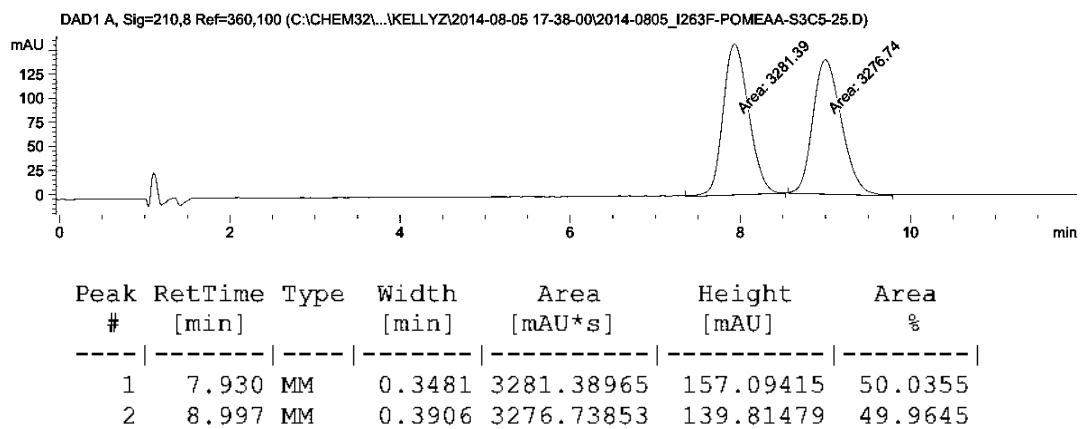


***N*-(2-hydroxy-2-(4-methoxyphenyl)ethyl)-4-methylbenzenesulfonamide (16).** SFC (Chiralpak AS-H column): 25% isopropanol in supercritical CO<sub>2</sub>, 210 nm.

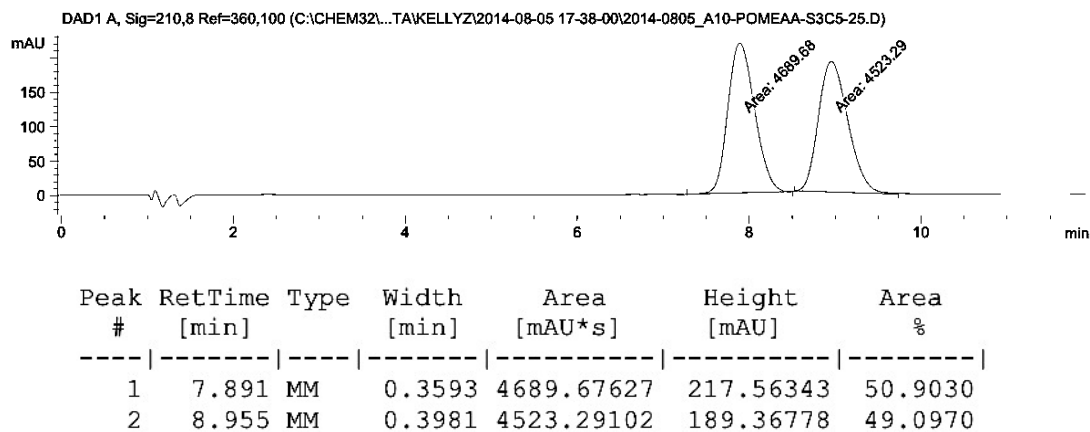
*Racemic reference compound*

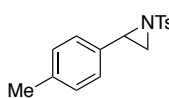


*P-I263F*



*P-I263F A328V L437V*

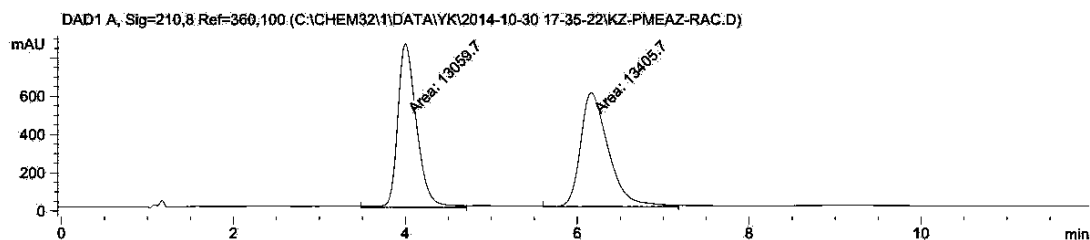




***N*-(*p*-Tolylsulfonyl)-2-(*p*-methylphenyl)aziridine (19).**

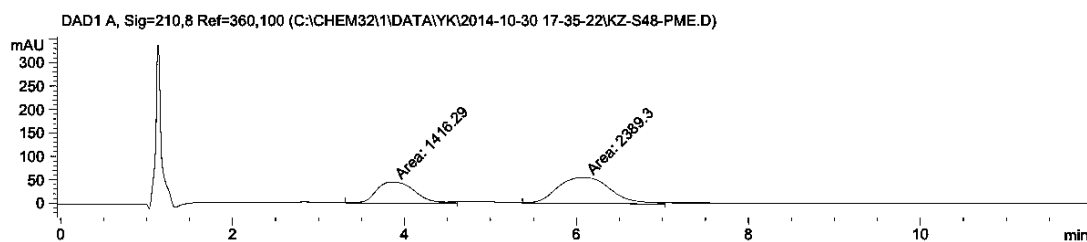
SFC (Chiralcel OB-H column): 20% isopropanol in supercritical CO<sub>2</sub>, 210 nm.

*Racemic reference compound*



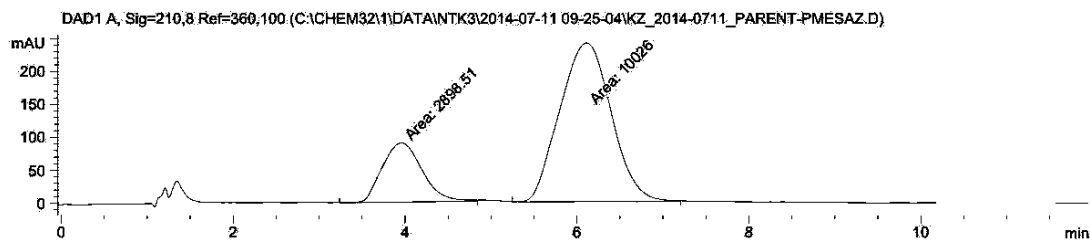
Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	4.000	MM	0.2545	1.30597e4	855.30292	49.3464
2	6.162	MM	0.3757	1.34057e4	594.68988	50.6536

*P411<sub>BM3</sub>-CIS T438S (P)*

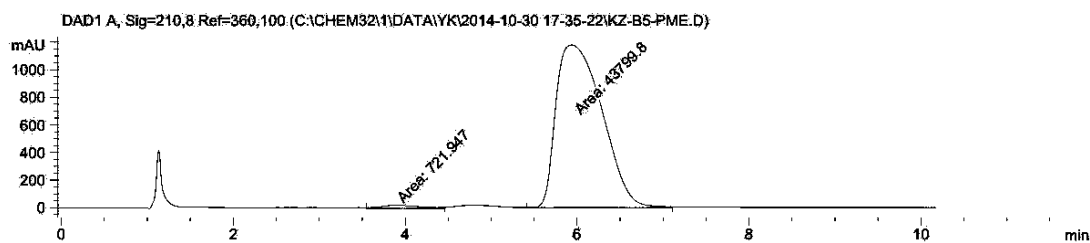


Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	3.857	MM	0.5209	1416.29333	45.31232	37.2161
2	6.084	MM	0.7149	2389.29932	55.70283	62.7839

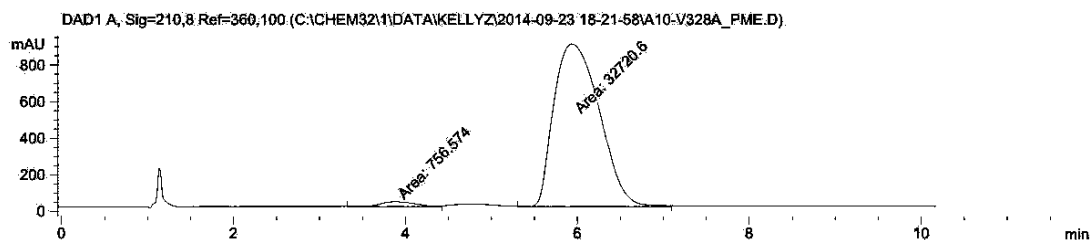
*P-I263F*



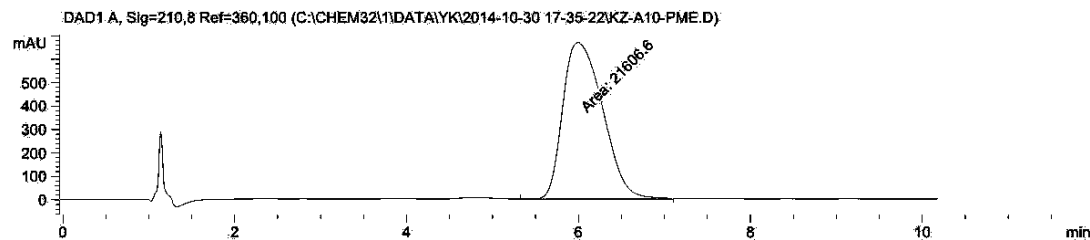
Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	3.960	MM	0.5373	2898.51050	89.90318	22.4264
2	6.110	MM	0.6936	1.00260e4	240.92017	77.5736

*P-I263F A328V*

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	3.869	MM	0.5473	721.94702	21.98478	1.6216
2	5.935	MM	0.6196	4.37998e4	1178.10986	98.3784

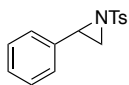
*P-I263F L437V*

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	3.876	MM	0.4836	756.57367	26.07187	2.2600
2	5.934	MM	0.6139	3.27206e4	888.33771	97.7400

*P-I263F A328V L437V*

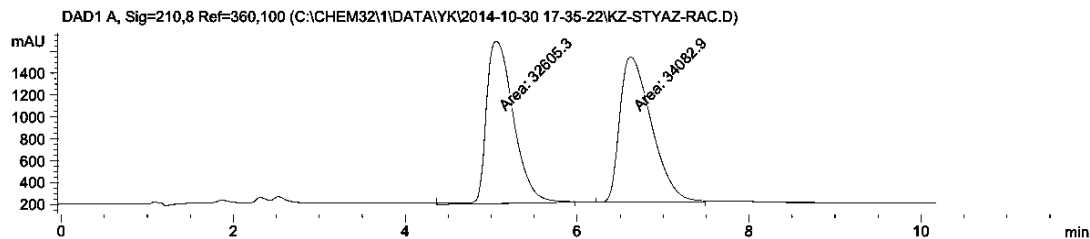
Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	5.992	MM	0.5394	2.16066e4	667.58752	100.0000





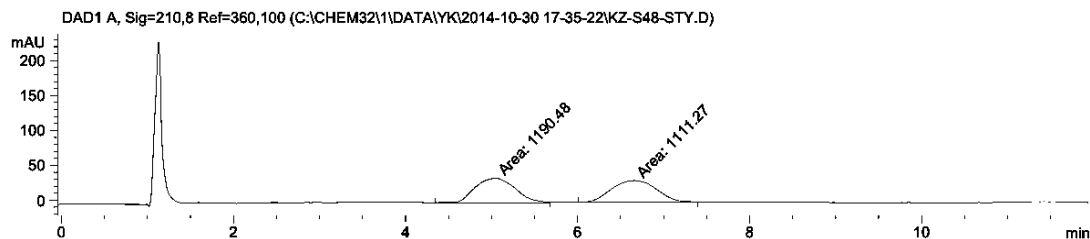
***N*-(*p*-Tolylsulfonyl)-2-phenylaziridine (21).** SFC (Chiralcel OB-H column): 15% isopropanol in supercritical CO<sub>2</sub>, 210 nm.

*Racemic reference compound*



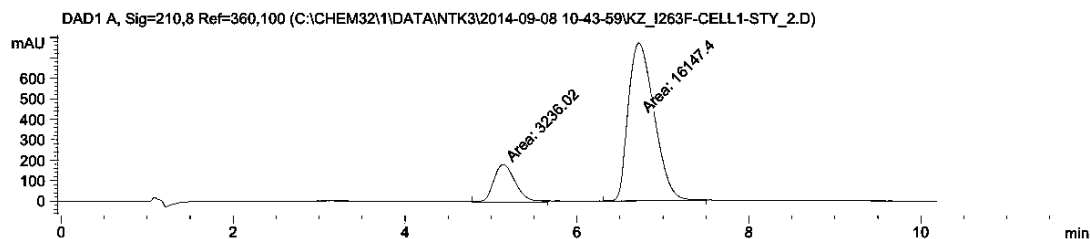
Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	5.057	MM	0.3659	3.26053e4	1485.14514	48.8922
2	6.622	MM	0.4271	3.40829e4	1329.99841	51.1078

*P411<sub>BM3</sub>-CIS T438S (P)*

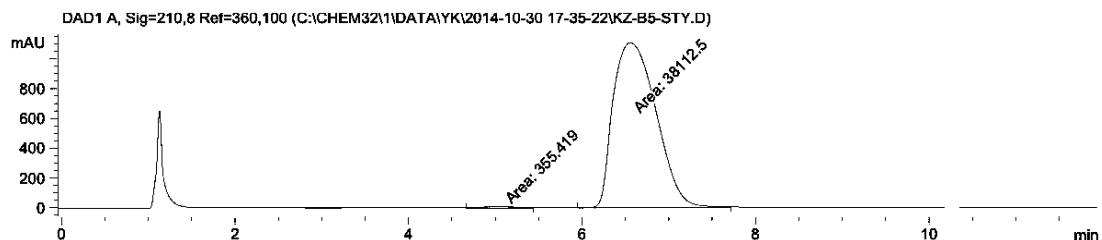


Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	5.040	MM	0.5540	1190.47607	35.81605	51.7206
2	6.651	MM	0.6016	1111.27014	30.78856	48.2794

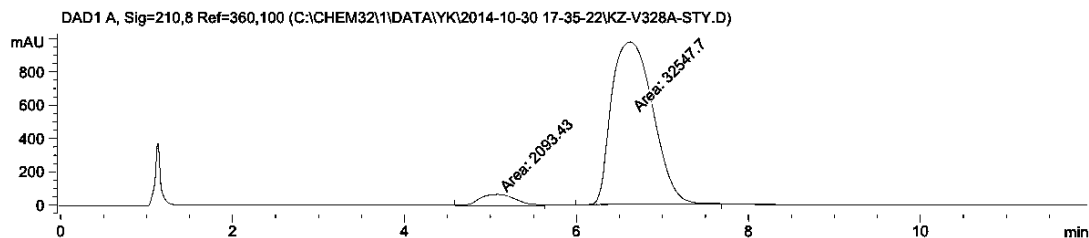
*P-I263F*



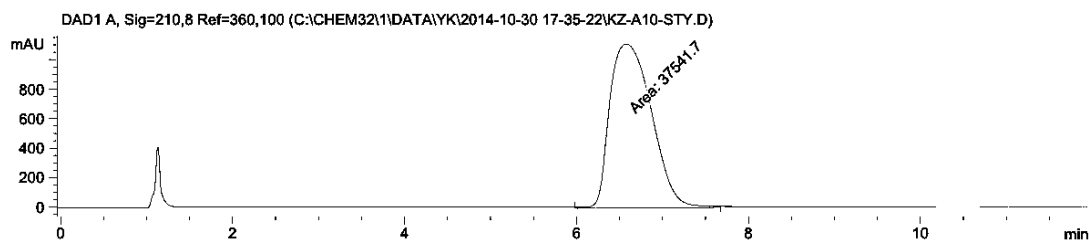
Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	5.140	MM	0.2953	3236.01807	182.66165	16.6948
2	6.720	MM	0.3487	1.61474e4	771.80231	83.3052

*P-I263F A328V*

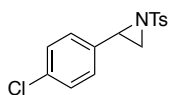
Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	5.077	MM	0.5269	355.41898	11.24200	0.9239
2	6.558	MM	0.5721	3.81125e4	1110.29065	99.0761

*P-I263F L437V*

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	5.078	MM	0.5042	2093.43091	69.20262	6.0432
2	6.624	MM	0.5566	3.25477e4	974.59937	93.9568

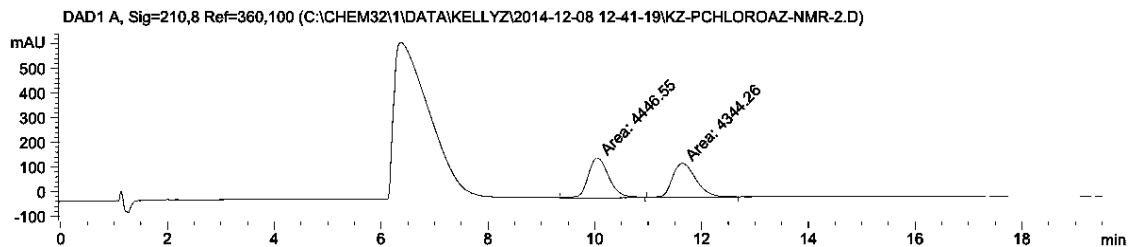
*P-I263F A328V L437V*

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	6.576	MM	0.5627	3.75417e4	1112.00940	100.0000



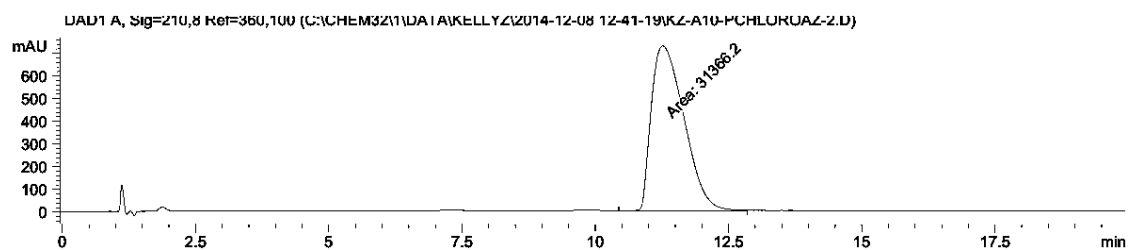
***N*-(*p*-Tolylsulfonyl)-2-(*p*-chlorophenyl)aziridine (22).** SFC (Chiralpak AS-H column): 8% isopropanol in supercritical CO<sub>2</sub>, 210 nm.

*Racemic reference compound*

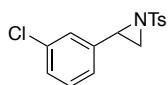


Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	10.048	MM	0.4545	4446.55078	163.06299	50.5818
2	11.641	MM	0.5222	4344.26123	138.65178	49.4182

*P-I263F A328V L437V*

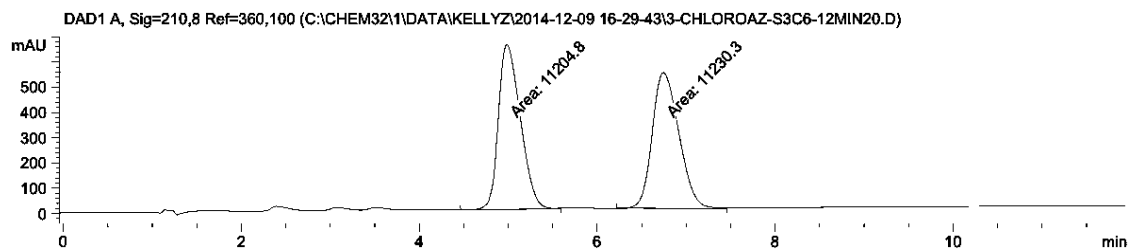


Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	11.265	MM	0.7168	3.13662e4	729.26678	100.0000



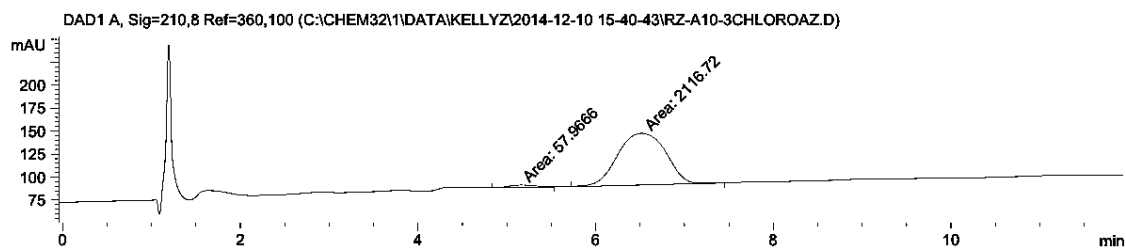
***N*-(*p*-Tolylsulfonyl)-2-(*m*-chlorophenyl)aziridine (23).** SFC (Chiralcel OB-H column): 20% isopropanol in supercritical CO<sub>2</sub>, 210 nm.

*Racemic reference compound*

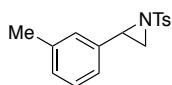


Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	4.985	MM	0.2862	1.12048e4	652.60254	49.9432
2	6.745	MM	0.3482	1.12303e4	537.54370	50.0568

*P-I263F A328V L437V*

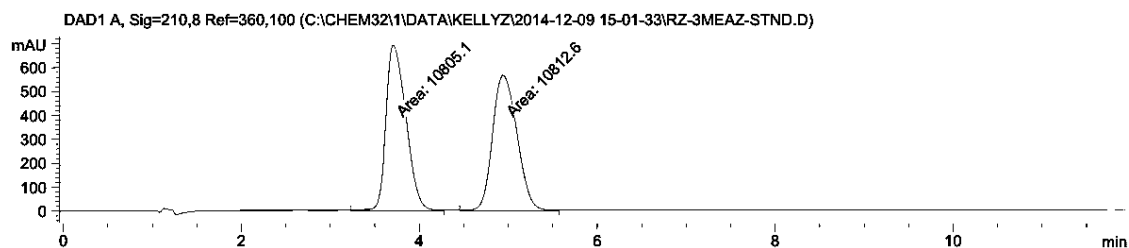


Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	5.145	MM	0.3705	57.96660	2.60781	2.6655
2	6.516	MM	0.6276	2116.71973	56.21136	97.3345



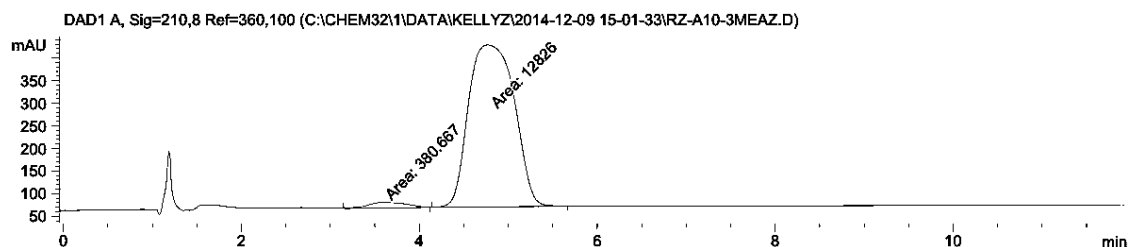
***N*-(*p*-Tolylsulfonyl)-2-(*m*-methylphenyl)aziridine (24).** SFC (Chiralcel OB-H column): 20% isopropanol in supercritical CO<sub>2</sub>, 210 nm.

*Racemic reference compound*

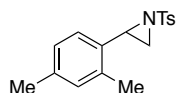


Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	3.708	MM	0.2602	1.08051e4	692.21600	49.9826
2	4.942	MM	0.3186	1.08126e4	565.55908	50.0174

*P-I263F A328V L437V*

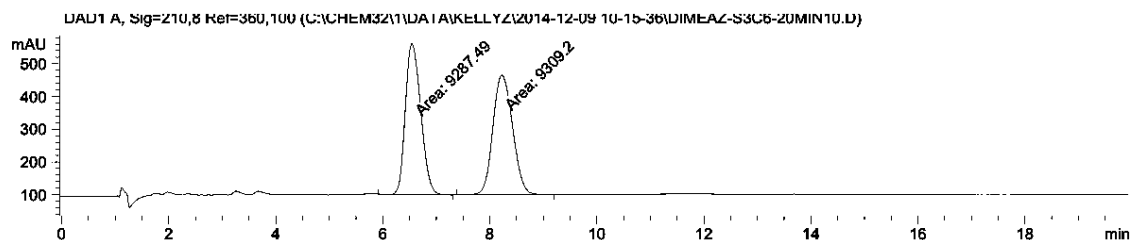


Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	3.579	MM	0.5097	380.66705	12.44634	2.8824
2	4.770	MM	0.5976	1.28260e4	357.67941	97.1176



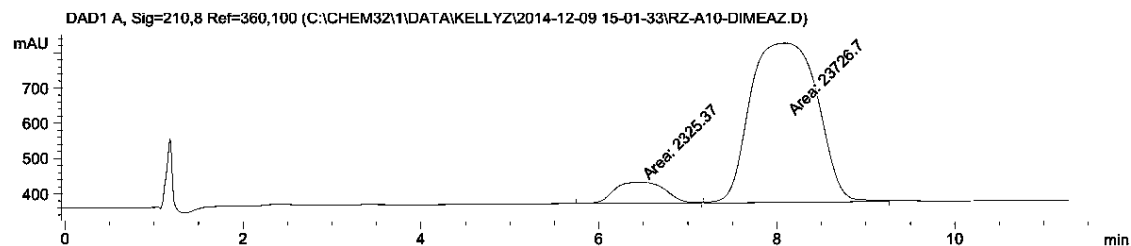
***N*-(*p*-Tolylsulfonyl)-2-(2,4-dimethylphenyl)aziridine (25).** SFC  
(Chiralcel OB-H column): 10% isopropanol in supercritical CO<sub>2</sub>, 210  
nm.

*Racemic reference compound*

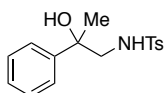


Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	6.545	MM	0.3348	9287.48535	462.29697	49.9416
2	8.226	MM	0.4250	9309.19922	365.07251	50.0584

*P-I263F A328V L437V*

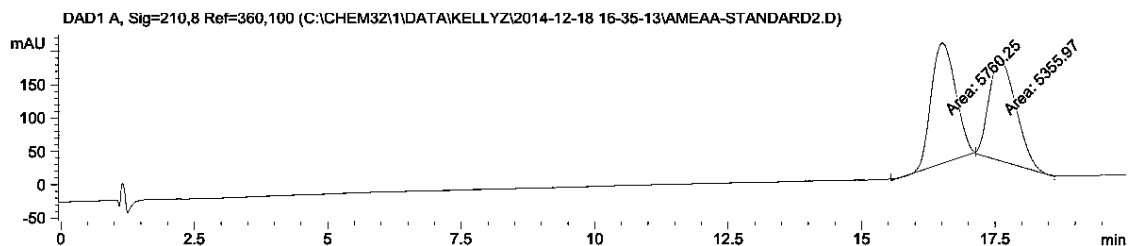


Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	6.455	MM	0.6568	2325.36816	59.01160	8.9258
2	8.093	MM	0.8759	2.37267e4	451.45123	91.0742



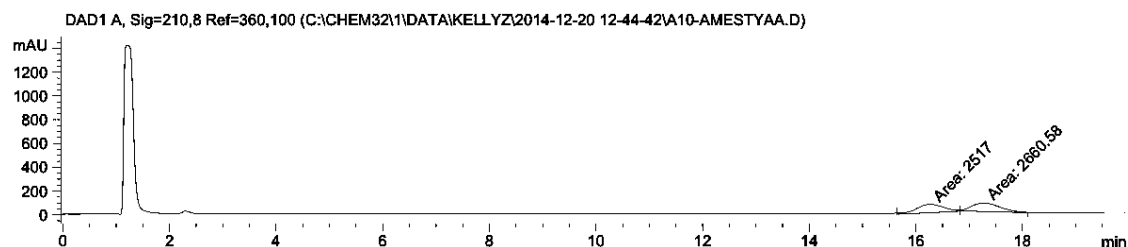
***N*-(2-hydroxy-2-phenylpropyl)-4-methylbenzenesulfonamide (26).** SFC (Chiralpak AS-H column): 15% isopropanol in supercritical CO<sub>2</sub>, 210 nm.

*Racemic reference compound*

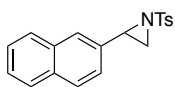


Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	16.511	MM	0.5311	5760.24756	180.76260	51.8184
2	17.607	MM	0.5928	5355.97217	150.58235	48.1816

*P-I263F A328V L437V*

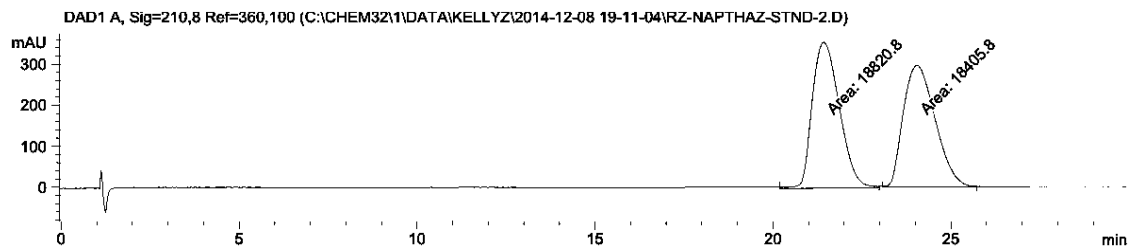


Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	16.265	MM	0.5942	2517.00049	70.60035	48.6135
2	17.276	MM	0.6129	2660.57593	72.35230	51.3865



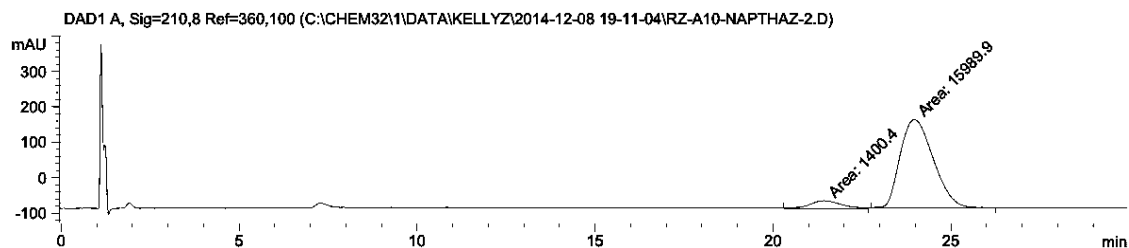
***N*-(*p*-Tolylsulfonyl)-2-(2-naphthyl)aziridine (27).** SFC (Chiralpak AS-H column): 8% isopropanol in supercritical CO<sub>2</sub>, 210 nm.

*Racemic reference compound*



Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	21.429	MM	0.8822	1.88208e4	355.55453	50.5574
2	24.049	MM	1.0350	1.84058e4	296.38309	49.4426

*P-I263F A328V L437V*

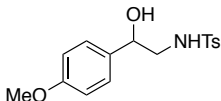
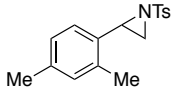
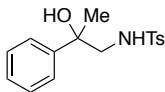


Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	21.465	MM	1.0441	1400.40186	22.35329	8.0528
2	23.976	MM	1.0731	1.59899e4	248.33498	91.9472



### A.7 $^1\text{H}$ and $^{13}\text{C}$ NMR Spectra

$^1\text{H}$  and  $^{13}\text{C}$  NMR spectra of the following compounds can be found in the Supporting Information of the published paper:

Compound	Spectra
	$^1\text{H}$ (400 MHz, $\text{CDCl}_3$ ), $^{13}\text{C}$ (400 MHz, $\text{CDCl}_3$ )
	$^1\text{H}$ (400 MHz, $\text{CDCl}_3$ ), $^{13}\text{C}$ (400 MHz, $\text{CDCl}_3$ )
	$^1\text{H}$ (400 MHz, $\text{CDCl}_3$ ), $^{13}\text{C}$ (400 MHz, $\text{CDCl}_3$ )

## A.8 References and Notes

1. Farwell, C. C. et al. Enantioselective imidation of sulfides via enzyme-catalyzed intermolecular nitrogen-atom transfer. *J. Am. Chem. Soc.* **2014**, *136*, 8766–8771.
2. Heel, T.; McIntosh, J.A.; Dodani, S.C.; Meyerowitz, J.T.; Arnold, F.H. Non-natural olefin cyclopropanation catalyzed by diverse cytochrome P450s and other hemoproteins. *ChemBioChem.* **2014**, *15*, 2556.
3. Srinivas, B.; Kumar, V.P.; Sridhar, R.; Surendra, K.; Nageswar, Y.V.D; Rao, K.R. J. Regioselective nucleophilic opening of epoxides and aziridines under neutral conditions in the presence of  $\beta$ -cyclodextrin in water. *J. Mol. Catal. A: Chem.* **2007**, *261*, 1–5.
4. Ando, T.; Kano, D.; Minakata, S.; Ryu, I; Komatsu, M. Iodine-catalyzed aziridination of alkenes using chloramine-T as a nitrogen source. *Tetrahedron* **1998**, *54*, 13485–13494.
5. Evans, D. A.; Faul, M. M.; Bilodeau, M. T. Development of the copper-catalyzed olefin aziridination reaction. *J. Am. Chem. Soc.* **1994**, *116*, 2742–2753.
6. Huang, C. Y.; Doyle, A. G. Nickel-catalyzed Negishi alkylations of styrenyl aziridines. *J. Am. Chem. Soc.* **2012**, *134*, 9541–9544.
7. Craig II, R.A.; O'Connor, N.R.; Goldberg, A.F.G.; Stoltz, B.M. Stereoselective lewis acid mediated (3+2) cycloadditions of *N*-H- and *N*-sulfonylaziridines with heterocumulenes. *Chem. Eur. J.* **2014**, *20*, 4806 – 4813.
8. Gao, G.Y.; Harden, J.D.; Zhang, X.P. Cobalt-catalyzed efficient aziridination of alkenes. *Org. Lett.* **2005**, *7*, 3191–3193.
9. Takeda Y.; Ikeda, Y.; Kuroda, A.; Tanaka, S.; Minakata, S. Pd/NHC-Catalyzed enantiospecific and regioselective Suzuki–Miyaura arylation of 2-arylaziridines: Synthesis of enantioenriched 2-arylphenethylamine derivatives. *J. Am. Chem. Soc.* **2014**, *136*, 8544–8547.
10. Alonso, D.A.; Andersson, P.G. Deprotection of sulfonyl aziridines. *J. Org. Chem.* **1998**, *63*, 9455–9461.
11. Wang, X.; Ding, K. One-pot enantioselective aziridination of olefins catalyzed by a copper(I) complex of a novel diimine ligand by using  $\text{PhI}(\text{OAc})_2$  and sulfonamide as nitrene precursors. *Chem. Eur. J.* **2006**, *12*, 4568–4575.

## INTERMOLECULAR BENZYLIC C–H AMINATION CATALYZED BY ENGINEERED CYTOCHROME P450 ENZYMES

Material for this chapter appears in Prier, C. K.<sup>†</sup>; **Zhang, R. K.**<sup>†</sup>; Buller, A. R.; Brinkmann-Chen, S.; Arnold, F. H. “Enantioselective, intermolecular benzylic C–H amination catalysed by an engineered iron-haem enzyme,” *Nature Chemistry* **2017**, 9, 629–634. DOI: 10.1038/nchem.2783. (<sup>†</sup>Denotes equal contribution) This work was performed in collaboration with all authors. Reprinted with permission from Nature Publishing Group.

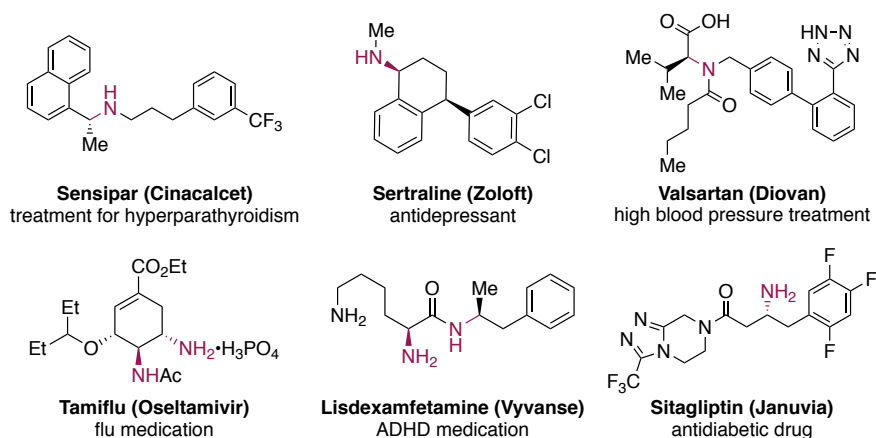
### 3.1 Abstract

C–H bonds are ubiquitous structural units of organic molecules. While these bonds were traditionally considered to be chemically inert, the recent emergence of methods for C–H functionalization promises to transform the way synthetic chemistry is performed. Due to the prevalence of nitrogen in bioactive molecules, the intermolecular amination of C–H bonds represents a particularly desirable transformation. Here we report the directed evolution of an iron-containing enzyme catalyst, based on a cytochrome P450 monooxygenase, for the highly enantioselective, intermolecular amination of benzylic C–H bonds. The biocatalyst is capable of up to 1,300 turnovers, exhibits excellent enantioselectivities, and provides access to valuable benzylic amines. Although iron complexes are generally poor catalysts for C–H amination, the protein framework enables this transformation and does so with exquisite control of enantioselectivity. This work demonstrates that a renewable protein catalyst based on iron, the most abundant transition metal in the earth's crust, can solve an outstanding challenge in synthetic chemistry.

### 3.2 Introduction

Chiral amine compounds are valuable intermediates in the synthesis of pharmaceuticals and other target molecules. The chiral amine moiety is widely present in top-selling drugs, including Cinacalcet (Sensipar), Sertraline (Zoloft), and Valsartan (Diovan), as well as many others (Figure 3-1).<sup>1</sup> Methods for their preparation typically require the addition of a stoichiometric auxiliary (e.g. chiral *tert*-butanesulfinamide<sup>2</sup>),

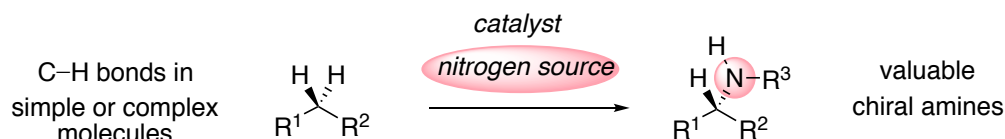
kinetic resolution (e.g. by organocatalysis<sup>3</sup> or biocatalysis<sup>4</sup>), or the use of a transition metal catalyst with chiral ligands (e.g. asymmetric hydrogenation<sup>5</sup>, hydroamination<sup>6</sup>, or C–H functionalization<sup>7</sup>). Due to the demonstrated importance of the chiral amine functionality, continued research efforts to develop efficient and sustainable methods for their preparation, such as the use of base metal catalysts and biocatalysts, will make a huge impact on synthesis.



**Figure 3-1.** Examples of chiral amine containing pharmaceutical agents.

We were intrigued by the direct transformation of an  $sp^3$  hybridized C–H bond to a C–N bond, which can be accomplished via the generation and transfer of a metal nitrene.<sup>7–9</sup> Many of the methods that have been developed for this transformation are intramolecular,<sup>10</sup> which requires that a nitrogen source already be present in the same molecule as the targeted C–H bond. An elegant yet more elusive transformation is the intermolecular C–H amination of starting materials which do not contain a required functionality; such a reaction provides a dramatically simplified and convergent disconnection for the synthesis of amines (Figure 3-2). Recognizing the potential of this transformation, the development of innovative catalytic systems for intermolecular C–H amination have been pursued. While many of these methods are racemic, enantioselective intermolecular C–H amination via nitrene transfer has been developed using rhodium,<sup>11–13</sup> ruthenium,<sup>14,15</sup> and manganese<sup>15,16</sup> catalysts. Highly diastereoselective C–H amination

using a chiral nitrogen source has also been achieved using a rhodium catalyst.<sup>17</sup> These systems represent major advances in C–H functionalization; none, however, offers the combination of high catalytic efficiency (turnover), high enantioselectivity across diverse substrates, and ready access to the chiral catalyst.



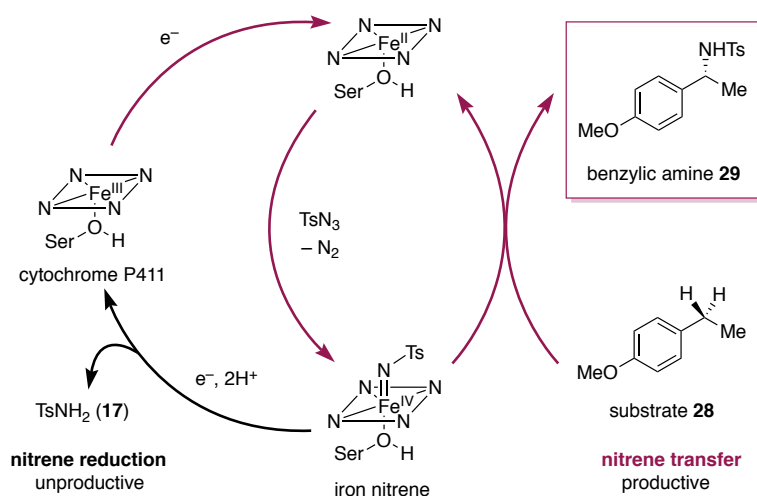
- Challenging transformation with no highly enantioselective, earth-abundant catalysts
- Direct C–H amination is not a natural function of any known enzyme

**Figure 3-2.** Intermolecular C–H amination, a simplifying transformation for chiral amine synthesis. Intermolecular C–H amination enables direct and convergent functionalization in which a substrate and a nitrogen atom source are brought together in a single step. In principle, any C–H bond in the substrate is a potential site of functionalization.

Curiously, a process for the direct transformation of a  $sp^3$  C–H bond to C–N bond is not known to be a natural function of any known enzyme. Biosynthetic routes to amines instead appear to exclusively rely on functional group manipulation of pre-oxidized substrates, and biocatalytic multi-enzyme cascades have been engineered for formal C–H amination using this general approach.<sup>18,19</sup> We and others, however, have demonstrated that heme proteins can catalyze nitrene transfer, including for intramolecular C–H amination,<sup>20–24</sup> under appropriate reaction conditions. In particular, variants of cytochrome P450<sub>BM3</sub> from *Bacillus megaterium* that feature a serine axial ligand to the heme iron in place of the wild-type cysteine ligand, which we term cytochrome P411s,<sup>25</sup> are proficient catalysts for these reactions.

Building on Chapter 2 and other work in the Arnold group, here we describe the directed evolution of P411 enzymes for intermolecular benzylic C–H amination. Similar to the proposed aziridination mechanism, we envisioned a catalytic cycle for intermolecular amination as shown in Figure 3-3. First, reduction of the ferric state of the heme cofactor gives the catalytically active ferrous state. Reaction with a nitrene source, here tosyl azide (TsN<sub>3</sub>), then provides the putative iron nitrene intermediate. Subsequent

reaction of this intermediate with an alkane substrate such as 4-ethylanisole (**28**) would deliver the C–H amination product **29** and regenerate the ferrous state of the catalyst. A competing process observed in P411-catalyzed nitrene transfer is the reduction of the iron nitrene, generating the undesired by-product *p*-toluenesulfonamide (TsNH<sub>2</sub>, **17**). While heme protein catalysts have been identified for intramolecular C–H amination, in such cases the targeted C–H bond has an implicit proximity advantage to the nitrene intermediate due to the intramolecular arrangement. In contrast, a catalyst for the desired intermolecular amination must bind a separate alkane substrate and promote its amination faster than the transient nitrene species is consumed in the reductive side reaction.



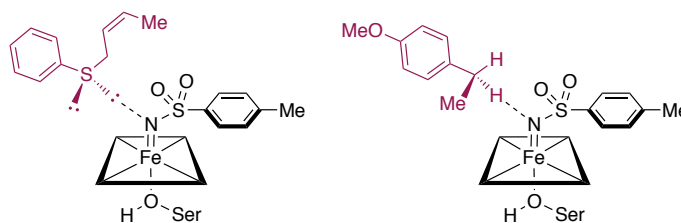
**Figure 3-3.** Proposed mechanism of intermolecular C–H amination catalyzed by a cytochrome P411 enzyme. Reaction of tosyl azide with the ferrous heme cofactor generates an enzyme-bound iron nitrene intermediate. This nitrene then inserts into a C–H bond in a second substrate, delivering a benzylic amine product. The nitrogen atoms in a plane represent the enzyme's heme cofactor. Ts = 4-toluenesulfonyl; Ser = serine. Notes: to date, the protonation state of the axial serine ligand in P411 enzymes is not known; a formal oxidation state of the iron nitrene intermediate is shown, but may not accurately reflect its electronic structure.

### 3.3 Results and Discussion

#### 3.3.1 Reaction Discovery and Directed Evolution

In initial studies, we found that almost all cytochrome P411s reported to date display no activity toward intermolecular C–H amination of substrate **28** with tosyl azide, instead delivering exclusively sulfonamide **17** (Table B-1 in Appendix B). This includes catalysts that were developed for ostensibly similar reactions such as intermolecular aziridination (Chapter 2) and intramolecular C–H amination with arylsulfonyl azides.<sup>21,22</sup> Other heme proteins, as well as the heme cofactor alone (iron protoporphyrin IX) were similarly inactive (Table B-2 in Appendix B).

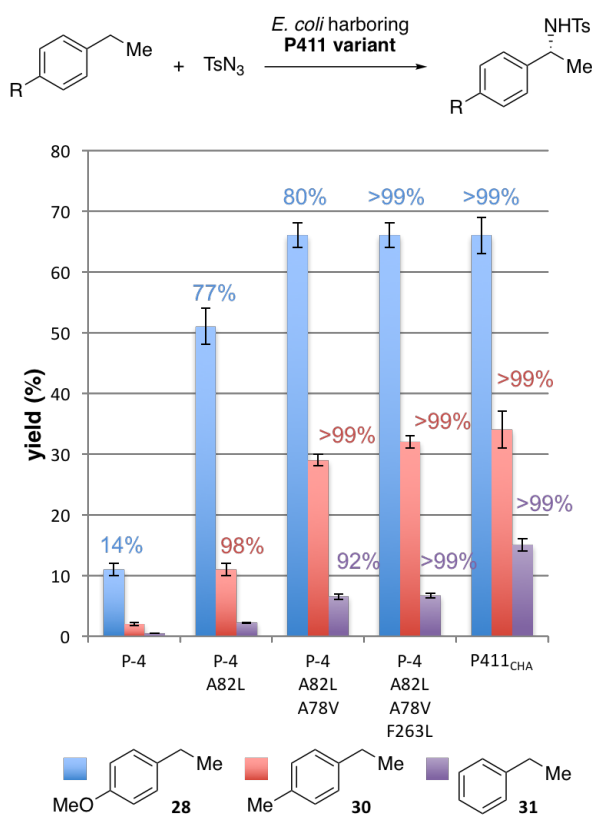
Remarkably, Dr. Christopher Prier identified a P411 variant, P-4, previously engineered for the imidation of allylic sulfides,<sup>26</sup> that showed some activity toward the desired C–H amination reaction. Under anaerobic conditions, variant P-4, which differs from wild-type P450<sub>BM3</sub> by 17 mutations, catalyzes the amination of 4-ethylanisole (**28**) at the benzylic carbon in 11% yield and 14% ee. Testing other variants created for the sulfide imidation reaction identified P-4 A82L, with an additional A82L mutation, which provides greater than 4-fold improvement in yield (to 51%) and delivers the product **4** in 77% ee (Table B-3 in Appendix B). Although it is not clear why evolution for sulfide imidation engendered activity for C–H amination, we can speculate that this may be due to similarities between the molecular geometries at the reactive positions of the two substrates (Figure 3-4). Like 4-ethylanisole, which is tetrahedral at the benzylic carbon, the sulfide substrate is also expected to be near tetrahedral at the reactive sulfur atom. In contrast, the olefin moiety which undergoes the aziridination reaction is planar.



**Figure 3-4.** Comparison of sulfide imidation (left) and C–H amination (right) substrate geometries. Substrates for both reactions have tetrahedral molecular geometry at the reactive position. The plane represents the enzyme's heme cofactor; nitrogen atoms have been omitted for clarity.

We next performed sequential rounds of site-saturation mutagenesis of selected residues in P-4 A82L and screening to improve C–H amination activity and enantioselectivity. We targeted sites in the heme domain that were either previously shown to impact the activity and/or selectivity of P450s, residues that are highly conserved in P450s, or residues that were already mutated in P-4 relative to wild-type P450<sub>BM3</sub> (see Table B-5 in Appendix B for details).<sup>27</sup> The libraries were screened for enhanced C–H amination activity, and potential hits were subsequently evaluated for activity across a set of three substrates with different electronic demands for C–H amination: 4-ethylanisole (**28**), 4-ethyltoluene (**30**), and ethylbenzene (**31**) (Figure 3-5). Two mutations, A78V and F263L, improve activity and selectivity on all three substrates, yielding a catalyst that delivers the benzylic amine products as single enantiomers (>99% ee). A final mutation, E267D, is neutral with respect to the amination of 4-ethylanisole but provides a two-fold improvement in the amination of ethylbenzene (**31**, to 15% yield), a less electronically activated substrate. We call this final variant P411<sub>CHA</sub>, for "cytochrome P411 C–H aminase." Under the conditions employed for evolution, this catalyst generates the benzylic amine **29** in 66% yield, with >99% ee and 1,000 turnovers in whole *E. coli* cells.



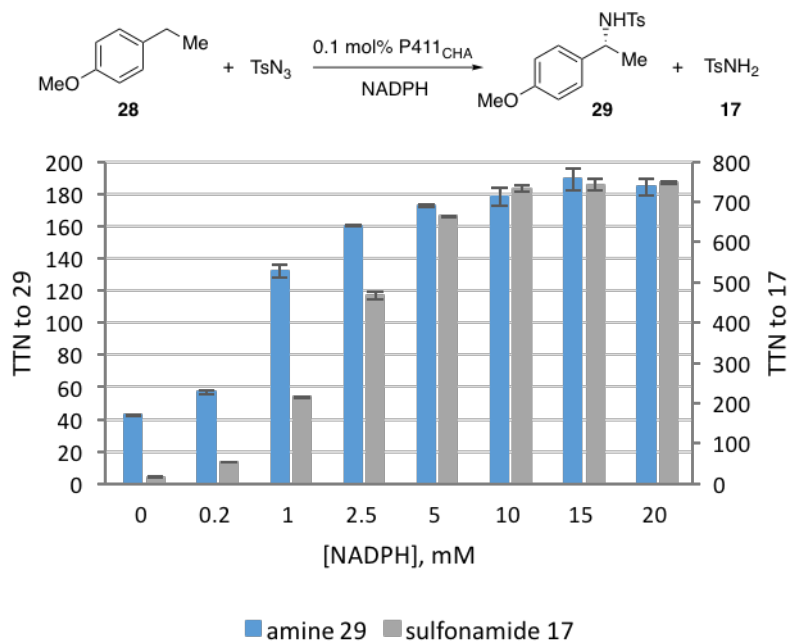


**Figure 3-5.** Evolution of a cytochrome P411 catalyst for enantioselective C–H amination. Directed evolution via sequential rounds of site-saturation mutagenesis and screening, improved both the yield and enantioselectivity of P411-catalyzed C–H amination. Reactions were performed using *E. coli* cells expressing the P411 variant, resuspended to  $OD_{600} = 30$ , with 5 mM alkane substrate and 5 mM tosyl azide, under anaerobic conditions. Results are the average of experiments performed with duplicate cell cultures, each used to perform duplicate chemical reactions (four reactions total). Bars represent yield; numbers above bars represent enantiomeric excess (ee); both are color-coded to match the substrate (blue, **28**; red, **30**; purple, **31**). Error bars correspond to one standard deviation. P-4 gives predominantly the *S* enantiomer in the amination of **28**; all other variant/substrate combinations give predominantly the *R* enantiomer. Data are tabulated in Table B-6 of Appendix B.

### 3.3.2 *In Vitro* Studies

With a variant optimized for intermolecular C–H amination of ethylbenzenes in hand, we characterized the activity of the enzyme under *in vitro* conditions. Using purified P411<sub>CHA</sub>, we investigated the effect of NADPH concentration, previously shown to be a suitable reductant for P411-catalyzed olefin aziridination reactions, on C–H amination activity (Figure 3-6). Surprisingly, in the absence of external reductant, a small amount of benzylic amine product was detected (ca. ~40 TTN). Addition of sub-stoichiometric amounts of NADPH (1 mM) significantly increased turnover to amine product (130 TTN to **29**) while stoichiometric and super-stoichiometric concentrations of NADPH (5 mM or above) gave a further improvement (up to ~180 TTN). The amount of sulfonamide **17** detected in reaction mixtures also increased with addition of up to 10 mM NADPH. These results support the hypothesis that the ferrous enzyme is active for the amination chemistry.

Increasing NADPH concentration may be beneficial to activity through reduction of ferric enzymes created during the course of catalysis (from production of sulfonamide **17**, for example, Figure 3-3), back to the catalytically active ferrous state.

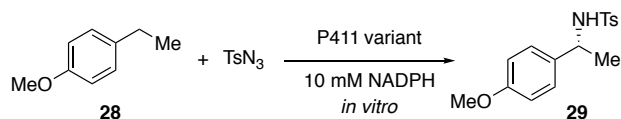


**Figure 3-6.** Effect of NADPH concentration on C–H amination activity of P411<sub>CHA</sub> *in vitro*. Reactions performed with 5  $\mu\text{M}$  P411<sub>CHA</sub> purified protein, NADPH (variable), 5 mM 4-ethylanisole, and 5 mM tosyl azide as described in the Experimental Methods (Section 3.5). Bars represent the average of duplicate reactions; error bars represent one standard deviation.

Employing 10 mM NADPH as the reductant, the four enzymes of the C–H amination lineage were evaluated *in vitro* using model substrates 4-ethylanisole and tosyl azide (Table 3-1). Notably, the amination activity of the purified proteins (up to 190 TTN to **29**) is much lower than in *E. coli* cells (1,000 TTN with P411<sub>CHA</sub>); further investigation of differences between *in vitro* and in whole cell conditions was not pursued. Increasing the enzyme concentration from 5  $\mu\text{M}$  to 20  $\mu\text{M}$  resulted in a higher yield of amination product **29** (from 15% to 47%), though with some loss to TTN (entries 4 and 5, Table 3-1). Characterization of initial rates of reaction with these enzymes demonstrate that the accumulated mutations increase the rate of the amination reaction: indeed, the initial rate of C–H amination was

enhanced by greater than 8-fold for reaction by P411<sub>CHA</sub> compared to starting variant P-4 (Table 3-2).

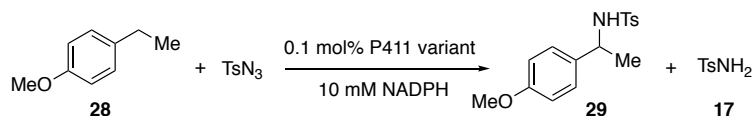
**Table 3-1.** C–H amination of 4-ethylanisole performed with purified P411 variants.<sup>a</sup>



Entry	Catalyst	Yield, TTN, ee of 29
1	P-4 (5 $\mu$ M)	2.2% yield, 22 TTN, –4% ee ( <i>S</i> )
2	P-4 A82L (5 $\mu$ M)	9.7% yield, 97 TTN, 80% ee ( <i>R</i> )
3	P-4 A82L A78V F263L (5 $\mu$ M) <sup>b</sup>	19% yield, 190 TTN, >99% ee ( <i>R</i> )
4	P411 <sub>CHA</sub> (5 $\mu$ M)	15% yield, 150 TTN, >99% ee ( <i>R</i> )
5	P411 <sub>CHA</sub> (20 $\mu$ M)	47% yield, 120 TTN, >99% ee ( <i>R</i> )

<sup>a</sup>Reactions performed with 5 mM 4-ethylanisole, 5 mM tosyl azide, 10 mM NADPH and purified enzyme; results are the average of duplicate reactions. <sup>b</sup>Variant P-4 A82L A78V F263L outperforms P411<sub>CHA</sub> *in vitro*, demonstrating differences in enzyme activities under whole cell versus *in vitro* conditions.

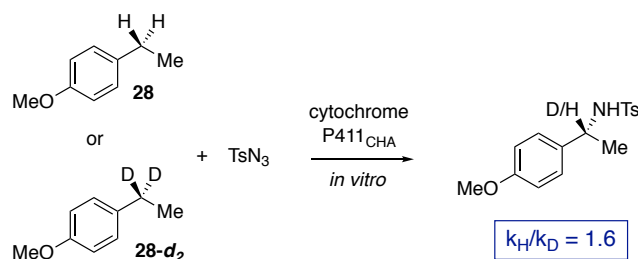
**Table 3-2.** Initial rates of intermolecular C–H amination and azide reduction.<sup>a</sup>



Variant	TOF 29	TOF TsNH <sub>2</sub> (17)
P-4	0.73 min <sup>–1</sup>	29.1 min <sup>–1</sup>
P-4 A82L	3.0 min <sup>–1</sup>	13.6 min <sup>–1</sup>
P-4 A82L A78V F263L <sup>b</sup>	7.3 min <sup>–1</sup>	12.2 min <sup>–1</sup>
P411 <sub>CHA</sub>	6.4 min <sup>–1</sup>	23.7 min <sup>–1</sup>

<sup>a</sup>Reactions performed with 5 mM 4-ethylanisole, 5 mM tosyl azide, 10 mM NADPH and 5  $\mu$ M purified enzyme; see Experimental Methods (Section 3.5) for details. Data used to calculate rates are presented in Figures B-1 and B-2 in Appendix B. <sup>b</sup>Variant P-4 A82L A78V F263L outperforms P411<sub>CHA</sub> *in vitro*, demonstrating differences in enzyme activities under whole cell versus *in vitro* conditions. TOF, turnover frequency.

Preliminary mechanistic studies were pursued to investigate the nature of the C–H insertion step. Independent rate measurements conducted with deuterated substrate **28-d<sub>2</sub>** show a kinetic isotope effect ( $k_H/k_D$ ) of 1.6 in the reaction catalyzed by P411<sub>CHA</sub>, suggesting C–H insertion effects the rate of the reaction (Figure 3-7).<sup>28</sup> A similar value (KIE = 1.9) was measured by Dubois and coworkers on Rh<sub>2</sub>(OAc)<sub>4</sub> mediated intramolecular C–H amination of sulfamates, although they employed an intramolecular competition experiment for this measurement so the values are not directly comparable.<sup>29</sup> KIE values of this magnitude have been associated with a concerted nitrene transfer step.<sup>30</sup> However, in the absence of additional computational or experimental studies, which were not conducted, we cannot conclude details regarding the mechanism.

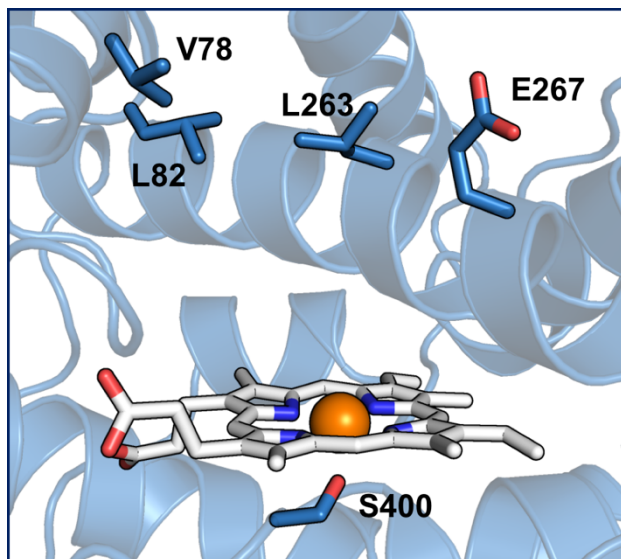


**Figure 3-7.** Kinetic isotope effect studies with P411<sub>CHA</sub>. The kinetic isotope effect for P411<sub>CHA</sub>-catalyzed C–H amination was determined from independent *in vitro* rate experiments; see Experimental Methods (Section 3.5) for details. Data used to calculate rates are presented in Figure B-3 in Appendix B

### 3.3.3 Structural Studies

We obtained an X-ray crystal structure of the penultimate variant in the intermolecular C–H amination lineage (P-4 A82L A78V F263L, heme domain only) at 1.70-Å resolution (see Table 3-6 in Section 3.5 Experimental Methods for crystallography analysis). *Note:* we only attempted crystallization of the heme domain because full length P450<sub>BM3</sub>, which contains three domains (heme domain, FMN binding domain, and FAD domain), has never been successfully crystallized. The beneficial mutations identified in this study are all located in helices that line the active site; two mutations are located in the enzyme's B' helix (A82L, A78V) and two are located in the I helix (F263L, E267D) (Figure 3-8). Residues in both of these helices are known to mediate substrate binding and/or impact

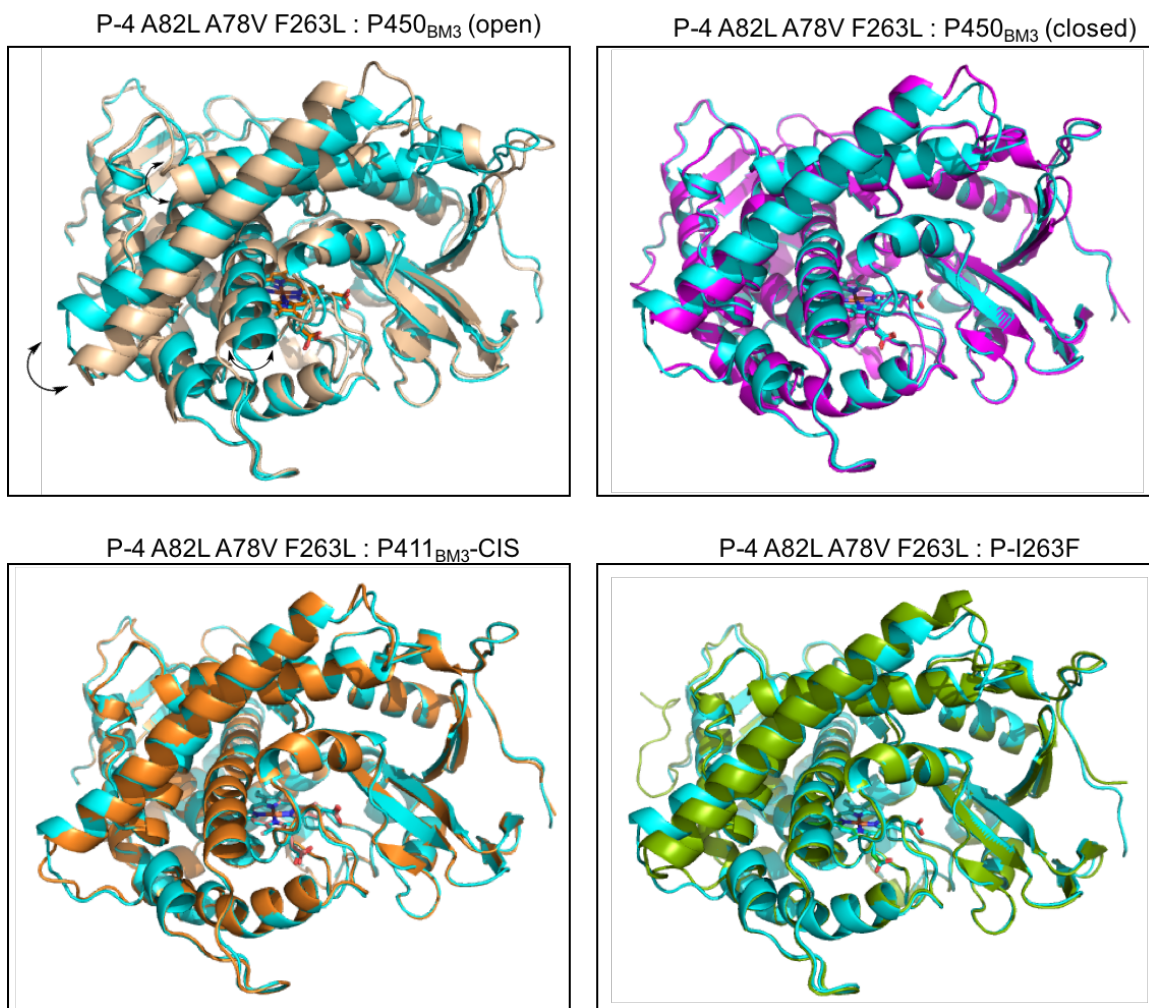
selectivity in P450-catalyzed oxygenation reactions.<sup>31–33</sup> Soaking experiments in which crystals of P-4 A82L A78V F263L heme domain were combined with 4-ethylanisole (**28**) or benzylic amine product (**29**) were unsuccessful at yielding a structure with substrate or product occupancy in the active site.



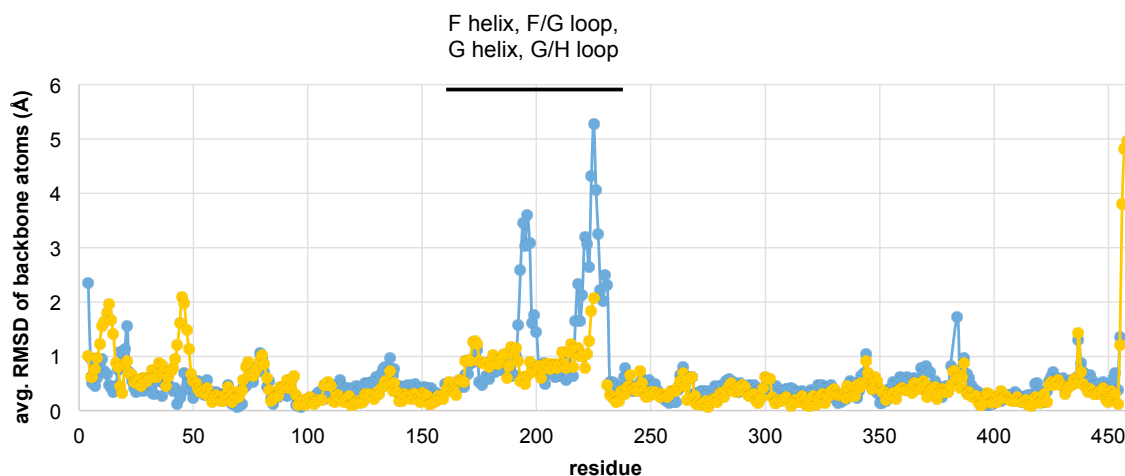
**Figure 3-8.** Active site view of the P-4 A82L A78V F263L crystal structure. The heme is shown in white and the iron atom in orange. Key active site residues are labeled and shown as sticks in blue. Residue S400 ligates the iron center; mutations at positions 78, 82, 263, and 267 enhance C–H amination activity and/or selectivity. All beneficial mutations identified in this study lie in the P411 active site on the distal face of the heme.

The structure of the evolved aminase adopts the P450 closed state typically induced by substrate binding, which was also observed with two earlier reported P411 structures (Figure 3-9).<sup>22,25</sup> Compared to the structure of P411 variant P-I263F,<sup>22</sup> which differs by only six mutations yet performs intermolecular C–H amination with trace activity (<1% yield), there are only minor movements of the protein scaffold. Globally, the largest differences in the position of protein backbone atoms are observed between the F and G helices and the F/G and G/H loops, which are known to be the most variable regions of P450s (Figure 3-10).<sup>32</sup> Local differences within the distal heme pocket were also observed; for example, active-site residue P329 adopts an endo conformation in P-4 A82L A78V F263L compared to exo in P-I263F. On average, the volume of the heme distal pocket is estimated to be reduced by approximately 10% in the evolved variant compared to P-I263F (calculations performed by CASTp server; see Chapter 3.5 Experimental Methods for details). The smaller active site potentially enforces productive substrate binding modes.

These observations suggest that the mutations introduced on the path to P411<sub>CHA</sub> exert local effects that modulate interactions with the azide and alkane substrates in the active site.



**Figure 3-9.** Heme domain protein alignments of cytochrome P411 variant P-4 A82L A78V F263L with wild type P450<sub>BM3</sub> and other P411<sub>BM3</sub> structures. Top panels show alignments of P411<sub>BM3</sub> P-4 A82L A78V F263L (cyan) with open (substrate free) form of wild type P450<sub>BM3</sub> (tan, PDB 2IJ2) and closed (substrate bound) form of wild-type P450<sub>BM3</sub> (magenta, PDB 1JPZ). Large movements of the I, F, and G helices are observed when comparing P-4 A82L A78V F263L with the open form of wild-type P450<sub>BM3</sub> (arrows); the position of these helices align more closely with the closed form of wild type P450<sub>BM3</sub>. Bottom panels show alignments of P-4 A82L A78V F263L (cyan) with previously obtained P411<sub>BM3</sub> structures, P411<sub>BM3</sub>-CIS (orange, PDB 4H23) and P-I263F (green, PDB 4WG2). No significant structural changes are observed. Protein alignments were carried about using the align tool of PyMOL (PyMOL Molecular Graphics System, Version 1.3, Schrodinger LLC).



**Figure 3-10.** Representative plots of root-mean-square deviation (RMSD) between the backbone atoms of P-4 A82L A78V F263L and P-I263F (PDB 4WG2). Blue: comparison of molecule A in P-4 A82L A78V F263L and molecule A in P-I263F. Yellow: comparison of molecule B of P-4 A82L A78V F263L and molecule C of P-I263F.

### 3.3.4 Substrate Scope

All substrate scope studies were carried out using P411<sub>CHA</sub> in *E. coli* cells. Subsequent experiments on modifying the reaction conditions revealed that P411<sub>CHA</sub> can support up to 1,300 turnovers (Table 3-3, entry 3), which exceeds the highest turnover number reported with any chiral transition metal complex to date for enantioselective, intermolecular C–H amination (85 TTN with a chiral manganese porphyrin).<sup>15</sup> Employing two equivalents of tosyl azide improved the yield of benzylic amine product: under these conditions, product **29** is obtained in 86% yield, with >99% ee and 670 turnovers (Table 3-3, entry 8).



**Table 3-3.** Optimization of the P411<sub>CHA</sub>-catalyzed amination of 4-ethylanisole.<sup>a b</sup>

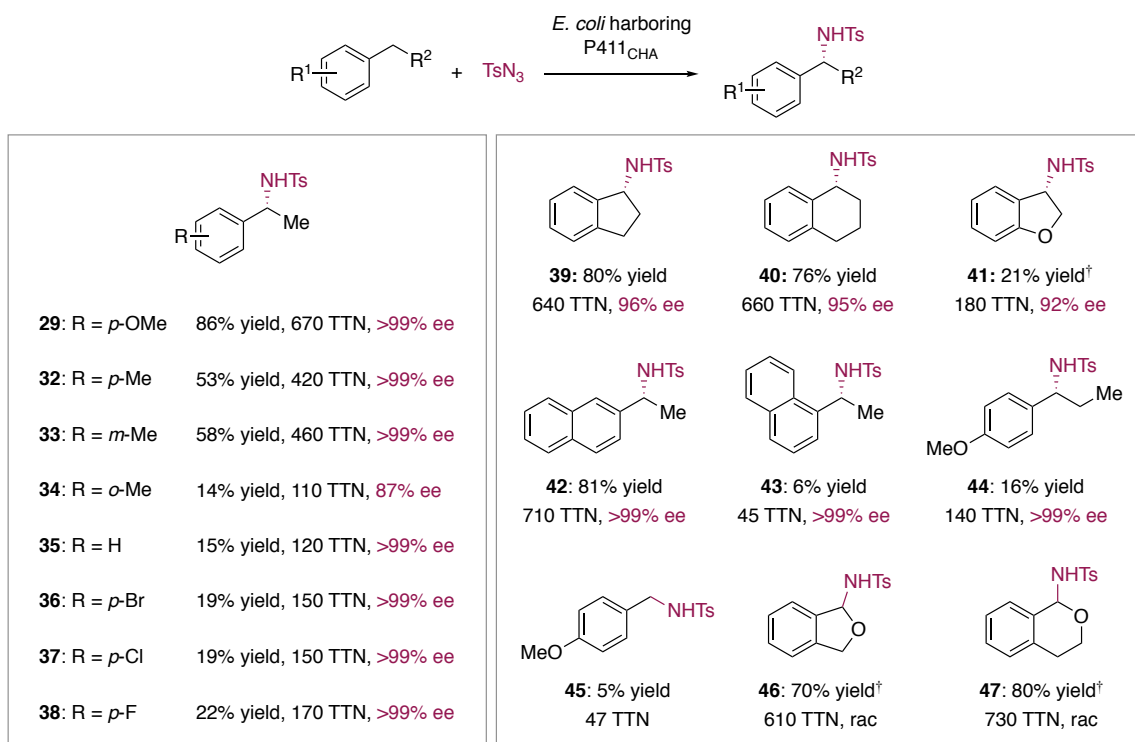
entry	mM 4-ethylanisole (3)	mM tosyl azide	yield (4)	TON
1	5	5	66%	1,000
2	7.5	7.5	51%	1,200
3	10	10	41%	1,300
4	15	15	26%	1,200
5	20	20	20%	1,200
6	2.5	2.5	56%	430
7	2.5	3.75	80%	630
8	2.5	5	86%	670
9	2.5	7.5	83%	640
10	2.5	10	78%	610

<sup>a</sup>Reactions were performed with *E. coli* cells expressing P411<sub>CHA</sub>, as in Figure 3-5. All reactions generate benzylic amine **29** in >99% ee. Blue, conditions employed during evolution; Red, turnover-optimized conditions; Purple, yield-optimized conditions, employed for evaluating the substrate scope. <sup>b</sup>These experiments were conducted by Dr. Chris Prier.

Scope studies, led by Dr. Chris Prier, were carried out using the yield-optimized conditions. Cytochrome P411<sub>CHA</sub> aminates a diverse set of arene-containing hydrocarbons. Substitution of the aromatic ring is tolerated at *para*, *meta*, and *ortho* positions (**32–34** in Figure 3-11). While electron-withdrawing functionality reduces the reactivity of the alkane substrate toward the metal nitrenoid, several halogenated ethylbenzenes are functionalized with >100 turnovers (**36–38**). The cyclic substrates indan and tetralin are excellent substrates (giving products **39** and **40**), while the related 2,3-dihydrobenzofuran displays reduced reactivity (product **41**). Larger alkylarenes such as ethylnaphthalenes and 4-propylanisole are still accommodated in the enzyme active site, undergoing amination with varying levels of efficiency (products **42–44**). Notably, the amination of 1-ethylnaphthalene delivers the nitrogen-containing fragment of the calcimimetic drug cinacalcet (Sensipar, Figure 3-1) with the correct absolute configuration (**43**).<sup>34</sup> Furthermore, although the activity is low, the amination of the methyl group of 4-methylanisole demonstrates that these catalysts are capable of functionalizing a primary benzylic C–H bond (product **45**). Interestingly, cumene and 4-methoxycumene are not aminated by P411<sub>CHA</sub> despite the presence of weak benzylic C–H bonds. Benzylic ethers



are suitable substrates for amination, but the products are isolated in racemic form (**46** and **47**). For the amination of isochroman (product **47**), analysis of crude reaction mixtures demonstrated that the C–H insertion event is moderately enantioselective (65% ee), but the product undergoes racemization during purification on silica gel (see Table B-7 in Appendix B). Notably, most of the alkane substrates evaluated undergo functionalization with excellent levels of enantioselectivity (>90% ee), a key advantage of this enzymatic approach over most reported metal catalysts.



**Figure 3-11.** Substrate scope of enzymatic intermolecular C–H amination. Reactions were performed in duplicate with *E. coli* cells expressing P411<sub>CHA</sub> at OD<sub>600</sub> = 30 (~3 μM enzyme), with 2.5 mM alkane substrate and 5.0 mM tosyl azide. Absolute configurations were assigned based on analogy to **29**, **32**, and **43**; rac, racemic. <sup>†</sup>Isolated yield from a reaction performed on 0.25-mmol scale (see Appendix B for details).

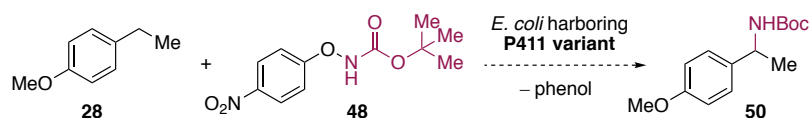
The enzymatic amination reaction can be performed on preparative scale: from a 0.25 mmol-scale biotransformation, the benzylic amine **29** was isolated in 78% yield (59.5 mg, 610 TTN, >99% ee). The tosyl group present in the C–H amination product may be

removed via treatment with samarium diiodide,<sup>35</sup> providing the corresponding primary amine with no erosion of enantiomeric excess (see Appendix B for details).

### 3.3.5 Extended Substrate Scope Studies

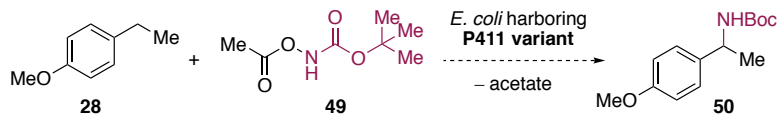
The harsh reductive conditions required for removal of the tosyl group limits the synthetic utility of this amination method. Carbamates, by contrast, are among the most widely used protecting groups for amines with many methods developed for their deprotection.<sup>36</sup> In addition, the carbamate moiety is present in bioactive molecules, including insecticides and therapeutics.<sup>37</sup> To test whether current heme protein variants could perform C–H amidation to install a carbamate group, a small set of proteins evolved for intermolecular C–H amination with tosyl azide were challenged with 4-ethylanisole (**28**) and nitrene precursor substrates **48** (Table 3-4) or **49** (Table 3-5). Notably, these nitrene precursor substrates are hydroxylamine derivatives; hydroxylamine reagents are preferred in order to avoid the use of low carbon organic azides (e.g. *tert*-butyl-oxycarbonyl azide (boc-azide), which has been implicated in explosions,<sup>38</sup> would be the corresponding azide reagent). Amination of *sp*<sup>3</sup>-hybridized C–H bonds using hydroxylamine reagents has been demonstrated with small molecule catalysts based on rhodium<sup>39, 40</sup> and copper.<sup>41</sup>

The tested heme protein variants showed no activity for the desired reactions. Enzyme reactions were extracted into organic solvent (1:1 mix of ethyl acetate and cyclohexane) and analyzed by gas chromatography-mass spectrometry (GC-MS). Comparison of traces from enzymatic reactions with separately synthesized reference compound **50** showed no product or signal close to background (trace). Additional experiments were not pursued.

**Table 3-4.** Initial tests for C–H amidation using reagent **48** with select cytochrome P411 variants.<sup>a</sup>

Variant	Mutations relative to P-4	Yield to <b>50</b>
P-4 A82L F263Y	A82L, F263Y	N.D.
P-4 A82L A78V F263L	A82L, A78V, F263L	N.D.
P-4 A82L A78V F263L A264X-A2	A82L, A78V, F263L, A264M	N.D.
P-4 A82L A78V F263L A74X-D8	A82L, A78V, F263L, A74C	N.D.
P-4 A82L A78V F263L A74X-B11	A82L, A78V, F263L, A74E	N.D.

<sup>a</sup>Reactions were conducted using *E. coli* cells expressing the appropriate P411 variant at OD<sub>600</sub> = 45, with 10 mM 4-ethylanisole and 10 mM reagent **48**. Cytochrome P411 variants were derived from the study described in this chapter. N.D., not detected. Amidation at the homobenzylic position was not observed using any of the tested enzymes.

**Table 3-5.** Initial tests for C–H amidation using reagent **49** with select cytochrome P411 variants.<sup>a</sup>

Variant	Mutations relative to P-4	Yield to <b>50</b>
P-4 A82L F263Y	A82L, F263Y	N.D.
P-4 A82L A78V F263L	A82L, A78V, F263L	N.D.
P-4 A82L A78V F263L A264X-A2	A82L, A78V, F263L, A264M	N.D.
P-4 A82L A78V F263L A74X-D8	A82L, A78V, F263L, A74C	trace
P-4 A82L A78V F263L A74X-B11	A82L, A78V, F263L, A74E	trace

<sup>a</sup>Reactions were conducted using *E. coli* cells expressing the appropriate P411 variant at OD<sub>600</sub> = 45, with 10 mM 4-ethylanisole and 10 mM reagent **49**. Cytochrome P411 variants were derived from the study described in this chapter. N.D., not detected; trace indicates a small signal, close to noise, was observed at the correct retention time. Amidation at the homobenzylic position was not observed using any of the tested enzymes.

### 3.4 Conclusion

Cytochrome P411<sub>CHA</sub> displays the ability to aminate benzylic C–H bonds intermolecularly in diverse structures with high selectivity, demonstrating that a renewable protein catalyst based on iron (the most abundant transition metal in the earth's crust)<sup>42</sup> can solve a long-standing challenge in synthetic chemistry. Biocatalysts for non-natural reactions have alternatively been created by introducing precious metals (such as iridium and rhodium) into proteins.<sup>43–46</sup> An artificial iridium metalloenzyme has been shown to perform nitrene transfer; this system is capable of up to ~300 turnovers in *intramolecular* C–H amination reactions.<sup>46</sup> That the P411 achieves intermolecular C–H amination with the native iron cofactor suggests that costly precious metals – and strategies for introduction of the non-native metal – are not necessary to achieve highly active biocatalysts for challenging non-natural reactions. The current work also describes an evolutionary pathway in which P411s evolved for a more readily accessible nitrene transfer reaction (sulfimination) picked up promiscuous activity toward a more challenging reaction (intermolecular C–H amination). This strategy of stepwise evolution through increasingly challenging activities may be generally useful for engineering enzymes for new activities not readily found by testing wild-type proteins. Finally, the ability to accelerate C–H insertion via mutation suggests that cytochrome P411<sub>CHA</sub> may be a platform for evolving catalysts for diverse C–H functionalization reactions currently inaccessible to chemical catalysis.

### 3.5 Experimental Methods

See Appendix B for supporting tables and figures, details of protein variants, characterization of compounds, details of calibration curves, assignment of absolute stereochemistry, and methods for determining enantioselectivity. Additional information can be found in the Supporting Information of the published paper.

### 3.5.1 General Information

Unless otherwise noted, all chemicals and reagents were obtained from commercial suppliers (Sigma-Aldrich, VWR, Alfa Aesar) and used without further purification. Silica gel chromatography was carried out using AMD Silica Gel 60, 230-400 mesh.  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra were recorded on a Varian Inova 300 MHz or 500 MHz, or Bruker Prodigy 400 MHz instrument, in  $\text{CDCl}_3$  and are referenced to residual protio solvent signals. Data for  $^1\text{H}$  NMR are reported as follows: chemical shift ( $\delta$  ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, p = pentet, m = multiplet, dd = doublet of doublets, dt = doublet of triplets, ddd = doublet of doublet of doublets), coupling constant (Hz), integration. Sonication was performed using a Qsonica Q500 sonicator. High-resolution mass spectra were obtained at the California Institute of Technology Mass Spectral Facility. Synthetic reactions were monitored using thin layer chromatography (Merck 60 gel plates) using a UV-lamp for visualization. Substrates were purchased from commercial suppliers. Tosyl azide was prepared according to de Nanteuil and Waser.<sup>47</sup>

### 3.5.2 Chromatography

Analytical high-performance liquid chromatography (HPLC) was carried out using an Agilent 1200 series instrument and a Kromasil 100 C18 column (4.6 x 50 mm, 5  $\mu\text{m}$ ) with water and acetonitrile as the mobile phase. Semi-preparative HPLC was performed using an Agilent XDB-C18 column (9.4 x 250 mm, 5  $\mu\text{m}$ ) with water and acetonitrile as the mobile phase. Liquid-chromatography-mass-spectrometry (LC-MS) was carried out using an Agilent 6140 series equipped with a C18 column with water (+0.1% acetic acid) and acetonitrile as mobile phases. Analytical chiral HPLC was conducted using a supercritical fluid chromatography (SFC) system with isopropanol and liquid  $\text{CO}_2$  as the mobile phase. Product enantiomers were separated using a Chiralpak AS column (4.6 x 150 mm, 5  $\mu\text{m}$ ) from Chiral Technologies Inc. Gas chromatography-mass spectrometry (GC-MS) analyses were carried out using a Shimadzu GCMS-QP2010SE system and J&W HP-5ms column (30 m x 0.25 mm, 0.25  $\mu\text{m}$  film).

### 3.5.3 Cloning and Site-Directed Mutagenesis

pET22b(+) was used as a cloning and expression vector for all enzymes described in this study. P450 and P411 enzymes described in this study were expressed with a C-terminal 6xHis-tag. Site-directed mutagenesis was performed using a modified QuikChange<sup>TM</sup> mutagenesis protocol.<sup>48</sup> The PCR products were gel purified, digested with DpnI, repaired using Gibson Mix<sup>TM</sup>,<sup>49</sup> and directly transformed into *E. coli* strain BL21 *E. coli* cells (Lucigen).

### 3.5.4 Determination of P411 Concentration

The concentration of P411 enzymes in whole cell experiments was determined from ferrous carbon monoxide binding difference spectra using the previously reported extinction coefficient for serine-ligated enzymes ( $\epsilon = 103,000 \text{ M}^{-1} \text{ cm}^{-1}$ ).<sup>50</sup> The concentration of purified P411 enzymes was determined by quantifying the amount of free heme present using the pyridine/hemochrome assay and the published extinction coefficient ( $\epsilon = 191,500 \text{ M}^{-1} \text{ cm}^{-1}$ ).<sup>51</sup>

### 3.5.5 Protein Expression and Amination Bioconversions Using Whole *E. coli* Cells

*E. coli* BL21 *E. coli* cells carrying a plasmid encoding a P411 variant were grown overnight in 5 mL Luria-Bertani medium with 0.1 mg/mL ampicillin (LB<sub>amp</sub>, 37 °C, 250 rpm). The preculture was used to inoculate 45 mL of Hyperbroth medium (supplemented with glucose nutrient mix according to package instructions) with 0.1 mg/mL ampicillin (HB<sub>amp</sub>) in a 125-mL Erlenmeyer flask; this culture was incubated at 37 °C, 230 rpm for 2 h. Cultures were then cooled on ice (20 min), and expression was induced with 0.5 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) and 1.0 mM 5-aminolevulinic acid (final concentrations). Expression was conducted at room temperature (23 °C), at 130 rpm, for 16–18 h. Cultures were then centrifuged (2,600 x g, 10 min, 4 °C), and the pellets were resuspended to an OD<sub>600</sub> of 30 (unless otherwise specified) in M9-N minimal media (no nitrogen). Aliquots of the cell suspension (4 mL) were used to determine the P411 expression level after lysis by sonication.

For amination bioconversions, the cells ( $OD_{600}$  of 30 in M9-N media) were degassed by sparging with argon in sealed 6-mL crimp vials for at least 40 minutes. Separately, a glucose solution (250 mM in M9-N) was degassed by sparging with argon for at least 10 minutes. An oxygen depletion system (20  $\mu$ L of a stock solution containing 14,000 U/mL catalase and 1,000 U/mL glucose oxidase in 0.1 M KPi, pH 8.0) was added to 2-mL crimp vials. All solutions were uncapped and transferred into an anaerobic chamber. Resuspended cells (320  $\mu$ L) were added to the vials, followed by glucose (40  $\mu$ L, 250 mM in M9-N), alkane substrate (10  $\mu$ L of a DMSO stock), and tosyl azide (10  $\mu$ L of a DMSO stock). Final concentrations were typically 2.5–5.0 mM alkane substrate, 5.0 mM tosyl azide, and 25 mM glucose; final reaction volume was 400  $\mu$ L. The vials were sealed, removed from the anaerobic chamber, and shaken at room temperature and 40 rpm for 16–20 h. The reactions were quenched by addition of acetonitrile (400  $\mu$ L) and internal standard (10  $\mu$ L of a DMSO stock). This mixture was then transferred to a microcentrifuge tube and centrifuged at 20,000  $\times$  g for 10 minutes. The supernatant was transferred to a vial and analyzed by HPLC; product concentrations were calculated using a calibration curve. For the determination of enantioselectivity, reaction mixtures were extracted with cyclohexane or cyclohexane : ethyl acetate (1:1, concentrated to dryness and re-suspended in an appropriate cyclohexane/ isopropanol mixture) and samples were analyzed by chiral SFC.

### 3.5.6 Reaction Screening in 96-well Plate Format.

Site-saturation libraries were generated employing the “22c-trick” method.<sup>48</sup> *E. coli* libraries were cultured in LB<sub>amp</sub> (300  $\mu$ L/well) at 37 °C, 220 rpm and 80% relative humidity overnight. Hyperbroth medium (1000  $\mu$ L/well, 0.1 mg/mL ampicillin) was inoculated with the preculture (50  $\mu$ L/well) and incubated at 37 °C, 220 rpm, 80% humidity for 3 h. The plates were cooled on ice for 30 minutes, and then expression was induced with 0.5 mM IPTG and 1.0 mM 5-aminolevulinic acid (final concentrations). Expression was conducted at 20 °C and 120 rpm for 24 h. The cells were pelleted (3,000  $\times$  g, 5 min, 4 °C) and resuspended in the oxygen depletion system (20  $\mu$ L/well). The 96-well plate was then transferred to an anaerobic chamber. In the anaerobic chamber, argon-sparged reaction

buffer (50 mM glucose in M9-N, 300  $\mu$ L/well) was added, followed by the alkane substrate (10  $\mu$ L/well, 200 mM in DMSO) and tosyl azide (10  $\mu$ L/well, 200 mM in DMSO). The plate was sealed with an aluminum foil, removed from the anaerobic chamber, and shaken at 40 rpm. After 16–20 h, the seal was removed and acetonitrile (400  $\mu$ L/well) and internal standard (10  $\mu$ L/well of a DMSO stock) were added. The wells were mixed, the plate was centrifuged (5,000 x g, 10 min), and the supernatant was filtered through an AcroPrep 96-well filter plate (0.2  $\mu$ m) into a shallow-well plate for HPLC analysis.

### 3.5.7 Protein Purification

*E. coli* BL21 *E. cloni* cells carrying a plasmid encoding a P411 variant were grown overnight in 25 mL LB<sub>amp</sub> (37 °C, 250 rpm). HB<sub>amp</sub> (630 mL, 0.1 mg/mL ampicillin) in a 1-L flask was inoculated with 25 mL of the preculture and incubated at 37 °C and 230 rpm for 2.5 h (to OD<sub>600</sub> ca. 1.8). Cultures were then cooled on ice (30 min) and induced with 0.5 mM IPTG and 1.0 mM 5-aminolevulinic acid (final concentrations). Expression was conducted at 23 °C, 130 rpm, for 16–20 h. Cultures were then centrifuged (5,000 x g, 8 min, 4 °C) and the cell pellets frozen at –20 °C. For purification, frozen cells from two such cultures were resuspended in buffer A (25 mM tris, 20 mM imidazole, 100 mM NaCl, pH 7.5, 4 mL/g of cell wet weight), loaded with hemin (1 mg/gram wet cell weight) and powdered DNaseI, and lysed by sonication. To pellet cell debris, lysates were centrifuged (20,000 x g, 20 min, 4 °C). Proteins were expressed in a construct containing a 6x-His tag and purified using a nickel NTA column (1 mL HisTrap HP, GE Healthcare, Piscataway, NJ) using an AKTA or AKTAexpress purifier FPLC system (GE healthcare). P411 enzymes were eluted with a linear gradient from 100% buffer A to 100% buffer B (25 mM tris, 300 mM imidazole, 100 mM NaCl, pH 7.5) over 10 column volumes.

Proteins used for crystallography were subjected to an additional ion-exchange purification step. For these proteins, fractions were pooled and subjected to three exchanges of anion exchange buffer A (25 mM tris-HCl, pH 7.5) using centrifugal spin filters (10 kDa molecular weight cut-off, Amicon Ultra, Merck Millipore). Next, the protein was loaded onto an anion exchange Q Sepharose column (HiTrap<sup>TM</sup> Q HP, GE



Healthcare) and purified using an AKTA or AKTAexpress purifier FPLC system (GE healthcare). The enzyme was eluted from the Q-column by running a gradient from 0 to 0.5 M NaCl over 10 column volumes. Fractions containing the enzyme were pooled, concentrated, and exchanged into storage buffer (25 mM tris-HCl, 25 mM NaCl, pH 7.5) as before. Subsequently, the concentrated protein was aliquoted, flash-frozen on powdered dry ice, and stored at  $-80^{\circ}\text{C}$ . Protein concentrations were determined via Bradford assay with a bovine serum albumin standard curve.

Proteins used for other purposes were eluted from the Ni-NTA column as described above, pooled, concentrated, and subjected to three exchanges of phosphate buffer (0.1 M KPi, pH 8.0) using centrifugal filters (10 kDa molecular weight cut-off, Amicon Ultra, Merck Millipore) to remove excess salt and imidazole. Concentrated proteins were aliquoted, flash-frozen on powdered dry ice, and stored at  $-80$  or  $-20^{\circ}\text{C}$ .

### **3.5.8 Amination Bioconversions Using Purified Protein**

Aliquots of phosphate buffer (260  $\mu\text{L}$  0.1 M potassium phosphate (KPi), pH 8.0) and NADPH (40  $\mu\text{L}$ , 100 mM), or multiples thereof, were combined in a 6-mL crimp vial and degassed by sparging with argon for at least 30 minutes. Separately, a glucose solution (250 mM in 0.1 M KPi, pH 8.0) was also degassed in the same manner. Crimp vials (2 mL) were each charged with the oxygen depletion system (20  $\mu\text{L}$  of a stock solution containing 14,000 U/mL catalase and 1,000 U/mL glucose oxidase in 0.1 M KPi, pH 8.0). After degassing was complete, all solutions, 2-mL crimp vials, and purified protein (100  $\mu\text{M}$  in 0.1 M KPi, pH 8.0), kept on ice, were brought into the anaerobic chamber. Glucose solution (40  $\mu\text{L}$ ), NADPH solution (300  $\mu\text{L}$ ), and purified protein (20  $\mu\text{L}$  of 100  $\mu\text{M}$  stock solution) were added to each 2-mL vial. Reaction vials were then charged with alkane substrate (10  $\mu\text{L}$ , 200 mM in DMSO) and tosyl azide (10  $\mu\text{L}$ , 200 mM in DMSO). Final concentrations were typically 5 mM alkane substrate, 5 mM tosyl azide, 10 mM NADPH, 25 mM glucose, and 5  $\mu\text{M}$  P411; final reaction volume was 400  $\mu\text{L}$ . Sodium dithionite (5 mM) was used as the reductant instead of NADPH for reactions with hemin and myoglobin. The vials were sealed, removed from the anaerobic chamber, and shaken at room temperature and 40 rpm

for 16–20 h. The reactions were quenched by addition of acetonitrile (400  $\mu$ L) and internal standard (10  $\mu$ L of a DMSO stock). This mixture was then transferred to a microcentrifuge tube and centrifuged at 20,000  $\times$  g for 10 minutes. The supernatant was transferred to a vial and analyzed by HPLC; product concentrations were calculated using a calibration curve. For the determination of enantioselectivity, reaction mixtures were extracted with cyclohexane or cyclohexane : ethyl acetate (1:1, concentrated to dryness and re-suspended in an appropriate cyclohexane/ isopropanol mixture) and samples were analyzed by SFC.

### 3.5.9 Determination of Initial Rates Using Purified Protein

Portions of phosphate buffer (260  $\mu$ L 0.1 M KPi, pH 8.0) and NADPH (40  $\mu$ L, 100 mM), or multiples thereof, were combined in a 6-mL crimp vial and degassed by sparging with argon for at least 30 minutes. Separately, glucose solution (250 mM in 0.1 M KPi, buffer pH 8.0) was also degassed in the same manner. In preparation, 2-mL crimp vials were each charged with a stir bar and the oxygen depletion system (20  $\mu$ L of a stock solution containing 14,000 U/mL catalase and 1,000 U/mL glucose oxidase in 0.1 M KPi, pH 8.0). After degassing was complete, all degassed solutions, prepared 2-mL crimp vials, and purified protein (100  $\mu$ M in 0.1 M KPi pH 8.0), kept on ice, were brought into the anaerobic chamber. Glucose solution (40  $\mu$ L), reaction solution (300  $\mu$ L), and purified protein (20  $\mu$ L of 100  $\mu$ M stock solution) were added to each 2-mL vial; the vials were placed on a stir plate and allowed to stir for 5 minutes. Reaction vials were then charged with alkane substrate (10  $\mu$ L, 200 mM in DMSO) and tosyl azide (10  $\mu$ L, 200 mM in DMSO). Final concentrations were 5 mM alkane substrate, 5 mM tosyl azide, 10 mM NADPH, 25 mM glucose, and 5  $\mu$ M P411; final reaction volumes were 400  $\mu$ L. Reactions were set up in duplicate and products quantified at 1-minute intervals by quenching with acetonitrile containing internal standard (410  $\mu$ L). This mixture was then removed from the anaerobic chamber, transferred to a microcentrifuge tube, and centrifuged at 20,000 $\times$ g for 10 minutes. The supernatant was transferred to a vial and analyzed by HPLC; product concentrations were calculated using a calibration curve. Rates of C–H amination and azide reduction were measured from the same reactions.

### 3.5.10 Crystallization, X-ray Data Collection, and Protein Structure Determination of P-4 A82L A78V F263L Heme Domain

The heme domain of variant P-4 A82L A78V F263L was crystallized by vapor diffusion. A 1:1 mixture of protein stock (22 mg/mL protein in 25 mM tris-HCl, 25 mM NaCl, pH 7.5 buffer) and mother liquor were combined in 24-well sitting drop plates (Hampton Research). The crystals were grown at room temperature over a span of 4–10 days. Crystals formed in 0.1 M bis-tris pH 4.0–5.0, 13% PEG 3350, 0.2 M NaHCOO. Data was collected on a crystal that formed under the following conditions: 0.1 M Bis-tris pH 5.0, 13% PEG 3350, 0.2 M NaHCOO. Crystals were cryo-protected by immersion into well solution with 25% glycerol before being flash-frozen in liquid N<sub>2</sub>. Diffraction data were collected on the Stanford Synchrotron Radiation Laboratory Beamline 12-2. Data was processed using XDS<sup>52</sup> in the space group P2<sub>1</sub>2<sub>1</sub>2 and scaled using AIMLESS<sup>53</sup> to 1.70-Å resolution.

The structure of P-4 A82L A78V F263L heme domain was solved by molecular replacement using PHASER, as implemented in CCP4, using the P411<sub>BM3</sub>-CIS T438S I263F structure (PDB ID: 4WG2) as the search model. Model building was performed in Coot and restrained refinement performed using Refmac5.<sup>54,55</sup> TLS operators were included in the last round of refinement.<sup>56</sup> Model quality was assessed with the MolProbity online server.<sup>57</sup> Crystallographic and model statistics are described in Table 3-6.

Soaking experiments with 4-ethylanisole (**28**) and benzylic amine **29** were performed as follows. Crystals of P-4 A82L A78V F263L heme domain were allowed to soak in well solution with 5–7.5 mM benzylic amine **29** or 5 mM 4-ethylanisole (**28**) and 5 to 15% DMSO as co-solvent for 24 h. Crystals survived these soaking conditions. Following, crystals were cryo-protected by immersion into well solution with 25% glycerol before being flash-frozen in liquid N<sub>2</sub>. Diffraction data were collected on the Stanford Synchrotron Radiation Laboratory Beamline 12-2. While some crystals which were soaked with substrate **28** and product **29** did diffract, occupancy of substrate or product was not observed in the active site of the protein.

**Table 3-6.** X-ray crystallography analysis of cytochrome P411 P-4 A82L A78V F263L heme domain (PDB: 5UCW).

<b>Data collection</b>	
Protein	Cytochrome P411 <sub>BM3</sub> P-4 A82L A78V F263L
PDB ID	5UCW
Beamline	SSRL 12.2 <sup>a</sup>
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	123.9, 127.0, 62.6
$\alpha$ , $\beta$ , $\gamma$ (°)	90.0, 90.0, 90.0
Wavelength	0.9795
Resolution (Å) (last bin A)	40.0 – 1.70
Last bin (Å)	1.73 – 1.70
<i>R</i> <sub>meas</sub> (%) <sup>b</sup>	5.5 (262.2)
<i>R</i> <sub>pim</sub> (%) <sup>b</sup>	1.5 (69.8)
CC <sub>1/2</sub> <sup>b</sup>	1.00 (0.593)
$\langle I / \sigma I \rangle^b$	23.9 (1.3)
Completeness (%) <sup>b</sup>	99.8 (100.0)
No. of observations	1,462,616
No. of unique reflections <sup>b</sup>	108,841 (5,345)
Redundancy <sup>b</sup>	13.4 (13.8)
<b>Refinement</b>	
Resolution (Å)	40.0 – 1.70
Final bin (Å)	1.744 – 1.700
No. of reflections	103,324
No. of reflections test set	5,400 (5.0%)
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub> <sup>b</sup>	18.9 / 22.3 (34.4 / 33.1)
Total no. atoms (non-hydrogen)	7,709
Average <i>B</i> -factor (Å <sup>2</sup> )	42.2
RMSD	
Bond lengths (Å)	0.012
Bond angles (°)	1.44
<b>Ramachandran plot<sup>c</sup></b>	
Favored (%)	97.3
Additionally allowed (%)	2.7
Outliers (%)	0.0
<b>MolProbity</b>	
Clashscore <sup>c</sup>	2.76 (99 <sup>th</sup> percentile)
Molprobity score <sup>c</sup>	1.29 (97 <sup>th</sup> percentile)

<sup>a</sup>SSRL, Stanford Synchrotron Radiation Lightsource. <sup>b</sup>Highest-resolution shell is shown in parentheses. <sup>c</sup>As determined by MolProbity.

*R*<sub>work</sub> is  $\|F_o - F_c\| / F_o$ , where *F*<sub>o</sub> is an observed amplitude and *F*<sub>c</sub> a calculated amplitude; *R*<sub>free</sub> is the same statistic calculated over a 5.0% subset of the data that has not been included during refinement.

### 3.5.11 Protein Cavity Volume Estimations

Volumes of the heme distal pocket were estimated by the CASTp server using a 1.4-Å probe.<sup>58</sup> Estimated volumes for P-4 A82L A78V F263L were 1489 and 1411 Å<sup>3</sup> for the two chains; estimated volumes for P-I263F (4WG2) were 1422, 1655, and 1737 Å<sup>3</sup> for the three chains.

## 3.5 References and Notes

1. Smith, D. T.; Delost, M. D.; Qureshi, H.; Njardarson, J. T. *Top 200 Pharmaceutical Products by Retail Sales in 2016*, **2018**, Available from: [https://njardarson.lab.arizona.edu/sites/njardarson.lab.arizona.edu/files/2016Top200PharmaceuticalRetailSalesPosterLowResV3\\_0.pdf](https://njardarson.lab.arizona.edu/sites/njardarson.lab.arizona.edu/files/2016Top200PharmaceuticalRetailSalesPosterLowResV3_0.pdf)
2. Robak, M. T.; Herbage, M. A.; Ellman, J. A. Synthesis and applications of *tert*-butanesulfinamide. *Chem. Rev.* **2010**, *110*, 3600–3740.
3. Kreituss, I.; Bode, J. W. Catalytic kinetic resolution of saturated *N*-heterocycles by enantioselective amidation with chiral hydroxamic acids. *Acc. Chem. Res.* **2016**, *49*, 2807–2821.
4. Kamal, A.; Azhar, M. A.; Krishnaji, T.; Malik, S.; Azeeza, S. Approaches based on enzyme mediated kinetic to dynamic kinetic resolutions: A versatile route for chiral intermediates. *Coord. Chem. Rev.* **2008**, *252*, 569–592.
5. Xie, J.-H.; Zhu, S.-F.; Zhou, Q.-L. Transition metal-catalyzed enantioselective hydrogenation of enamines and imines. *Chem. Rev.* **2011**, *111*, 1713–1760.
6. Hultsch, K. C. Transition metal-catalyzed asymmetric hydroamination of alkenes (AHA). *Adv. Synth. Catal.* **2005**, *347*, 367–391.
7. Collet, F.; Lescot, C.; Dauban, P. Catalytic C–H amination: the stereoselectivity issue. *Chem. Soc. Rev.* **2011**, *40*, 1926–1936.
8. Davies, H. M. L.; Manning, J. R. Catalytic C–H functionalization by metal carbenoid and nitrenoid insertion. *Nature* **2008**, *451*, 417–424.
9. Wu, W.-T.; Yang, Z.-P.; You, S.-L. Asymmetric C–H bond insertion reactions. In *Asymmetric Functionalization of C–H Bonds*, RSC Catalysis Series No. 25, **2015**, chapter 1.

10. Jeffrey, J. L.; Sarpong, R. Intramolecular C(sp<sup>3</sup>)-H amination. *Chem. Sci.* **2013**, *4*, 4092–4106.
11. Nägeli, I.; Baud, C.; Bernardinelli, G.; Jacquier, Y.; Moran, M.; Müller, P. Rhodium(II)-catalyzed CH insertions with {[ (4-nitrophenyl)sulfonyl]imino}phenyl- $\lambda^3$ -iodane. *Helv. Chim. Acta* **1997**, *80*, 1087–1105.
12. Yamawaki, M.; Tsutsui, H.; Kitagaki, S.; Anada, M.; Hashimoto, S. Dirhodium(II) tetrakis[*N*-tetrachlorophthaloyl-(*S*)-*tert*-leucinate]: a new chiral Rh(II) catalyst for enantioselective amidation of C–H bonds. *Tetrahedron Lett.* **2002**, *43*, 9561–9564.
13. Reddy, R. P.; Davies, H. M. L. Dirhodium tetracarboxylates derived from adamantylglycine as chiral catalysts for enantioselective C–H aminations. *Org. Lett.* **2006**, *8*, 5013–5016.
14. Nishioka, Y.; Uchida, T.; Katsuki, T. Enantio- and regioselective intermolecular benzylic and allylic C–H bond amination. *Angew. Chem. Int. Ed.* **2013**, *52*, 1739–1742.
15. Zhou, X.-G.; Yu, X.-Q.; Huang, J.-S.; Che, C.-M. Asymmetric amidation of saturated C–H bonds catalysed by chiral ruthenium and manganese porphyrins. *Chem. Commun.* **1999**, 2377–2378.
16. Kohmura, Y.; Katsuki, T. Mn(salen)-catalyzed enantioselective C–H amination. *Tetrahedron Lett.* **2001**, *42*, 3339–3342.
17. Liang, C.; Robert-Peillard, F.; Fruit, C.; Müller, P.; Dodd, R. H.; Dauban, P. Efficient diastereoselective intermolecular rhodium-catalyzed C–H amination. *Angew. Chem. Int. Ed.* **2006**, *45*, 4641–4644.
18. Both, P.; Busch, H.; Kelly, P. P.; Mutti, F. G.; Turner, N. J.; Flitsch, S. L. Whole-cell biocatalysts for stereoselective C–H amination reactions. *Angew. Chem. Int. Ed.* **2016**, *55*, 1511–1513.
19. Schrewe, M.; Ladkau, N.; Bühler, B.; Schmid, A. Direct terminal alkylamino-functionalization via multistep biocatalysis in one recombinant whole-cell catalyst. *Adv. Synth. Catal.* **2013**, *355*, 1693–1697.
20. Svastits, E. W.; Dawson, J. H.; Breslow, R.; Gellman, S. H. Functionalized nitrogen atom transfer catalyzed by cytochrome P-450. *J. Am. Chem. Soc.* **1985**, *107*, 6427–6428.
21. McIntosh, J. A.; Coelho, P. S.; Farwell, C. C.; Wang, Z. J.; Lewis, J. C.; Brown, T. R.; Arnold, F. H. Enantioselective intramolecular C–H amination catalyzed by engineered

- cytochrome P450 enzymes in vitro and in vivo. *Angew. Chem. Int. Ed.* **2013**, *52*, 9309–9312.
22. Hyster, T. K.; Farwell, C. C.; Buller, A. R.; McIntosh, J. A.; Arnold, F. H. Enzyme-controlled nitrogen-atom transfer enables regiodivergent C–H amination. *J. Am. Chem. Soc.* **2014**, *136*, 15505–15508.
23. Singh, R.; Bordeaux, M.; Fasan, R. P450-catalyzed intramolecular  $\text{sp}^3$  C–H amination with arylsulfonyl azide substrates. *ACS Catal.* **2014**, *4*, 546–552.
24. Bordeaux, M.; Singh, R.; Fasan, R. Intramolecular  $\text{C}(\text{sp}^3)$ –H amination of arylsulfonyl azides with engineered and artificial myoglobin-based catalysts. *Bioorg. Med. Chem.* **2014**, *22*, 5697–5704.
25. Coelho, P. S.; Wang, Z. J.; Ener, M. E.; Baril, S. A.; Kannan, A.; Arnold, F. H.; Brustad, E. M. A serine-substituted P450 catalyzes highly efficient carbene transfer to olefins *in vivo*. *Nature Chem. Biol.* **2013**, *9*, 485–487.
26. Prier, C. K.; Hyster, T. K.; Farwell, C. C.; Huang, A.; Arnold, F. H. Asymmetric enzymatic synthesis of allylic amines: A sigmatropic rearrangement strategy. *Angew. Chem. Int. Ed.* **2016**, *55*, 4711–4715.
27. Whitehouse, C. J. C.; Bell, S. G.; Wong, L.-L. P450<sub>BM3</sub> (CYP102A1): Connecting the dots. *Chem. Soc. Rev.* **2012**, *41*, 1218–1260.
28. Simmons, E. M.; Hartwig, J. F. On the interpretation of deuterium kinetic isotope effects in C–H bond functionalizations by transition-metal complexes. *Angew. Chem. Int. Ed.* **2012**, *51*, 3066–3072.
29. Fiori, K. W.; Espino, C. G.; Brodsky, B. H.; Du Bois, J. A mechanistic analysis of the Rh-catalyzed intramolecular C–H amination reaction. *Tetrahedron* **2009**, *65*, 3042–3051.
30. Varela-Álvarez, A. et al.  $\text{Rh}_2(\text{II,III})$  catalysts with chelating carboxylate and carboxamidate supports: Electronic structure and nitrene transfer reactivity. *J. Am. Chem. Soc.* **2016**, *138*, 2327–2341.
31. Poulos, T. L. Cytochrome P450. *Curr. Opin. Struct. Biol.* **1995**, *5*, 767–774.
32. Li, H.; Poulos, T. L. The structure of the cytochrome p450<sub>BM-3</sub> haem domain complexed with the fatty acid substrate, palmitoleic acid. *Nature Struct. Biol.* **1997**, *4*, 140–146.

- 
33. Roiban, G.-D.; Reetz, M. T. Expanding the toolbox of organic chemists: directed evolution of P450 monooxygenases as catalysts in regio- and stereoselective oxidative hydroxylation. *Chem. Commun.* **2015**, *51*, 2208–2224.
  34. Barniol-Xicota, M.; Leiva, R.; Escolano, C.; Vázquez, S. Synthesis of cinacalcet: An enantiopure active pharmaceutical ingredient (API). *Synthesis* **2016**, *48*, 783–803.
  35. Ankner, T.; Hilmersson, G. Instantaneous deprotection of tosylamides and esters with  $\text{SmI}_2$ /Amine/Water. *Org. Lett.* **2009**, *11*, 503–506.
  36. Wuts, P. G. M. *Greene's protective groups in organic synthesis*, 5th ed. Hoboken, N. J.: John Wiley & Sons, Inc. **2014**, chapter 7.
  37. Ghosh, A. K.; Brindisi, M. Organic carbamates in drug design and medicinal chemistry. *J. Med. Chem.* **2015**, *58*, 2895–2940.
  38. Feyen, P. Serious detonation during the preparation of *tert*-butyl-oxycarbonyl azide (boc-azide). *Angew. Chem. Int. Ed. Engl.* **1977**, *16*, 115.
  39. Lebel, H.; Huard, K.; Lectart, S. *N*-tosyloxycarbamates as a source of metal nitrenes: rhodium-catalyzed C–H insertion and aziridination reactions. *J. Am. Chem. Soc.* **2005**, *127*, 14198–14199.
  40. Reddy, R. P.; Davies, H. M. L. Dirhodium tetracarboxylates derived from adamantylglycine as chiral catalysts for enantioselective C–H aminations. *Org. Lett.* **2006**, *8*, 5013–5016.
  41. Wang, A.; Venditto, N. J.; Darcy, J. W.; Emmert, M. H. Nondirected, Cu-catalyzed  $\text{sp}^3$  C–H aminations with hydroxylamine-based amination reagents: catalytic and mechanistic studies. *Organometallics* **2017**, *36*, 1259–1268.
  42. Frey, P. A.; Reed, G. H. The ubiquity of iron. *ACS Chem. Biol.* **2012**, *7*, 1477–1481.
  43. Key, H. M.; Dydio, P.; Clark, D. S.; Hartwig, J. F. Abiological catalysis by artificial haem proteins containing noble metals in place of iron. *Nature* **2016**, *534*, 534–537.
  44. Srivastava, P.; Yang, H.; Ellis-Guardiola, K.; Lewis, J. C. Engineering a dirhodium artificial metalloenzyme for selective olefin cyclopropanation. *Nature Commun.* **2015**, *6*, 7789.
  45. Hyster, T. K.; Knörr, L.; Ward, T. R.; Rovis, T. Biotinylated Rh(III) complexes in engineered streptavidin for accelerated asymmetric C–H activation. *Science* **2012**, *338*, 500–503.



- 
46. Dydio, P.; Key, H. M.; Hayashi, H.; Clark, D. S.; Hartwig, J. F. Chemoselective, enzymatic C–H bond amination catalyzed by a cytochrome P450 containing an Ir(Me)-PIX cofactor. *J. Am. Chem. Soc.* **2017**, *139*, 1750–1753.
  47. de Nanteuil, F.; Waser, J. Catalytic [3+2] annulation of aminocyclopropanes for the enantiospecific synthesis of cyclopentylamines. *Angew. Chem. Int. Ed.* **2011**, *50*, 12075–12079.
  48. Kille, S.; Acevedo-Rocha, C. G.; Parra, L. P.; Zhang, Z.-G.; Opperman, D. J.; Reetz, M. T.; Acevedo J. P. Reducing codon redundancy and screening effort of combinatorial protein libraries created by saturation mutagenesis. *ACS Synth. Biol.* **2013**, *2*, 83–92.
  49. Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchinson III, C. A.; Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **2009**, *6*, 343–345.
  50. Vatsis, K. P.; Peng, H.-M.; Coon, M. J. Replacement of active-site cysteine-436 by serine converts cytochrome P450 2B4 into an NADPH oxidase with negligible monooxygenase activity. *J. Inorg. Biochem.* **2002**, *91*, 542–553.
  51. Fuhrhop, J.-H.; Smith, K. M. Laboratory methods in porphyrin and metalloporphyrin research. Elsevier, **1975**.
  52. Kabsch, W. XDS. *Acta Crystallographica Section D-Biological Crystallography* **2010**, *66*, 125–132.
  53. Evans, P. R.; Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallographica Section D-Biological Crystallography* **2013**, *69*, 1204–1214.
  54. Emsley, P.; Cowtan, K. *Coot*: model-building tools for molecular graphics. *Acta Crystallographica Section D-Biological Crystallography* **2004**, *60*, 2126–2132.
  55. Winn, M. D.; Murshudov, G. N.; Papiz, M. Z. Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol.* **2003**, *374*, 300–321.
  56. Painter, J.; Merritt, E. A. *TLSMD* web server for the generation of multi-group TLS models. *J. Appl. Cryst.* **2006**, *39*, 109–111.
  57. Chen, V. B.; Arendall III, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D-Biological Crystallography* **2010**, *66*, 12–21.

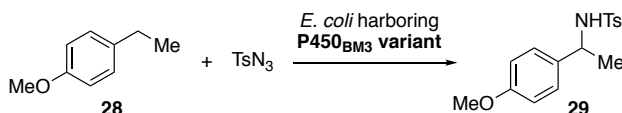
- 
58. Dundas, J.; Ouyang, Z.; Tseng, J.; Binkowski, A.; Turpaz, Y.; Liang, J. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucl. Acids Res.* **2006**, *34*, W116–W118.

## SUPPLEMENTARY INFORMATION FOR CHAPTER 3

Material for this chapter appears in Prier, C. K.<sup>†</sup>; **Zhang, R. K.**<sup>†</sup>; Buller, A. R.; Brinkmann-Chen, S.; Arnold, F. H. “Enantioselective, intermolecular benzylic C–H amination catalysed by an engineered iron-haem enzyme,” *Nature Chemistry* **2017**, 9, 629–634. DOI: 10.1038/nchem.2783. (<sup>†</sup>Denotes equal contribution) This work was performed in collaboration with all authors. Reprinted with permission from Nature Publishing Group.

## B.1 Supporting Tables B-1 through B-7

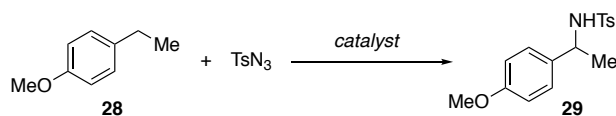
**Table B-1.** C–H amination of 4-ethylanisole with select variants of cytochrome P450<sub>BM3</sub>.<sup>a</sup>



Variant	Mutations relative to wild-type P450 <sub>BM3</sub>	Yield
pET22b(+) vector	N/A	N.D.
P450 <sub>BM3</sub>	None	N.D.
P411 <sub>BM3</sub>	C400S	N.D.
P411 <sub>BM3</sub> T268A	T268A, C400S	N.D.
P411 <sub>BM3</sub> -CIS T438S ("P")	V78A, F87V, P142S, T175I, A184V, S226R, H236Q, E252G, T268A, A290V, L353V, I366V, C400S, T438S, E442K	N.D.
P411 <sub>BM3</sub> -CIS A268T T438S	V78A, F87V, P142S, T175I, A184V, S226R, H236Q, E252G, A290V, L353V, I366V, C400S, T438S, E442K	N.D.
P411 <sub>BM3</sub> -H2-5-F10	L75A, V78A, F87V, P142S, T175I, A184V, S226R, H236Q, E252G, I263A, T268A, A290V, L353V, I366V, C400S, L437A, E442K	N.D.
P411 <sub>BM3</sub> -H2-A-10	L75A, V78A, F87V, P142S, T175I, L181A, A184V, S226R, H236Q, E252G, T268A, A290V, L353V, I366V, C400S, E442K	N.D.

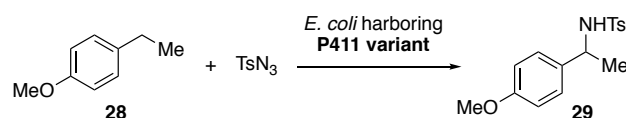
P411 <sub>BM3</sub> -H2-4-D4	L75A, V78A, F87V, P142S, T175I, M177A, L181A, A184V, S226R, H236Q, E252G, T268A, A290V, L353V, I366V, C400S, L437A, E442K	N.D.
P411 <sub>BM3</sub> T268A F87A <sup>b</sup>	F87A, T268A, C400S	N.D.
P-I263F <sup>b</sup>	V78A, F87V, P142S, T175I, A184V, S226R, H236Q, E252G, I263F, T268A, A290V, L353V, I366V, C400S, T438S, E442K	<1% yield
P-I263F A268T	P-I263F A268T	N.D.
P-I263F A328V	P-I263F A328V	<1% yield
P-I263F A328V L437V <sup>c</sup>	P-I263F A328V L437V	N.D.

<sup>a</sup>Reactions performed using *E. coli* cells expressing the appropriate protein at OD<sub>600</sub> = 30 with 5 mM of each substrate; results are the average of duplicate reactions. N.D., not detected. <sup>b</sup>Variants identified for regioselective intramolecular C–H amination.<sup>1</sup> <sup>c</sup>Variant identified for aziridination of styrenes (Chapter 2).<sup>2</sup>

**Table B-2.** C–H amination of 4-ethylanisole performed with hemin or select heme proteins.<sup>a</sup>

Catalyst	Yield
hemin (25 μM) <sup>b</sup>	N.D.
hemin (25 μM) + imidazole (1 mM) <sup>b</sup>	N.D.
hemin (25 μM) + bovine serum albumin (10 μM) <sup>b</sup>	N.D.
Myoglobin (Mb, equine heart, 10 μM) <sup>b</sup>	N.D.
Mb H64V V68A (sperm whale, 10 μM) <sup>b,c</sup>	N.D.
<i>Rhodothermus marinus</i> cytochrome c <sup>d</sup>	N.D.
<i>Hydrogenobacter thermophilus</i> cytochrome c <sup>d</sup>	N.D.
<i>Rhodopila globiformis</i> cytochrome c <sup>d</sup>	N.D.
<i>Rhodothermus marinus</i> cytochrome c V75T M100D M103E <sup>d,e</sup>	N.D.

<sup>a</sup>Reactions performed with 5 mM 4-ethylanisole and 5 mM tosyl azide; results are the average of duplicate reactions. N.D., none detected. <sup>b</sup>Performed *in vitro* with 5 mM sodium dithionite. <sup>c</sup>Variant identified for intramolecular C–H amination.<sup>3</sup> <sup>d</sup>Performed using *E. coli* cells at OD<sub>600</sub> = 30. <sup>e</sup>Variant identified for carbene Si–H insertion.<sup>4</sup>

**Table B-3.** C–H amination of 4-ethylanisole with P411 variants previously engineered for the imidation of allylic sulfides (ref. 5).<sup>a</sup>

Variant	Mutations relative to wild-type P450 <sub>BM3</sub>	Yield
P-I263F V87A	P-I263F V87A	2% yield
P-I263F V87A A328V	P-I263F V87A A328V	2% yield
P-4 <sup>b</sup>	P-I263F V87A A328V A268G	11% yield
P-I263F V87A A328V A82L	P-I263F V87A A328V A82L	4% yield
P-4 A82L <sup>b</sup>	P-I263F V87A A328V A268G A82L	51% yield
P-5	P-I263F V87A A328V A268G A82I	14% yield

<sup>a</sup>Reactions performed using *E. coli* cells expressing the appropriate protein at OD<sub>600</sub> = 30 with 5 mM of each substrate; results are the average of at least duplicate reactions. N.D., not detected. <sup>b</sup>This data also appears in Table B-6.

**Table B-4.** Mutations present in cytochrome P450<sub>BM3</sub> variants used in Chapter 3.

Variant	Mutations relative to wild type P450 <sub>BM3</sub>
P-4	V78A, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263F, T268G, A290V, A328V, L353V, I366V, C400S, T438S, E442K
P-4 A82L	V78A, A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263F, T268G, A290V, A328V, L353V, I366V, C400S, T438S, E442K
P-4 A82L A78V	A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263F, T268G, A290V, A328V, L353V, I366V, C400S, T438S, E442K
P-4 A82L A78V F263L	A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263L, T268G, A290V, A328V, L353V, I366V, C400S, T438S, E442K
P-4 A82L A78V F263L E267D (P411 <sub>CHA</sub> )	A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263L, E267D, T268G, A290V, A328V, L353V, I366V, C400S, T438S, E442K

**Table B-5.** Summary of directed evolution for intermolecular C–H amination.<sup>a</sup>

Gen.	Parent enzyme	Site-saturation libraries evaluated	Screening substrate	Mutation identified
1	P-4 A82L	A78X, L181X, F263X, T327X	4-ethylanisole	A78V
2	P-4 A82L A78V	F263X, T327X	4-ethylanisole	F263L
3a	P-4 A82L A78V F263L	A74X, A264X, F393X, G402X	4-ethylanisole	none
3b	P-4 A82L A78V F263L	R47X, S142X, V184X, G252X, E267X	4-ethyltoluene	E267D
4	P-4 A82L A78V F263L E267D (P411 <sub>CHA</sub> )	R47X, Y51X	4-ethyltoluene	none

<sup>a</sup>Some residues were saturated more than once, in different parent variants. Gen, generation.

Further information regarding residues:

A74, L181, F263, A264, E267, T327: Distal face of heme, active site residues

F393, G402: Axial face of heme

A78, S142, V184, G252: Already mutated in P-4 relative to wild type P450<sub>BM3</sub>

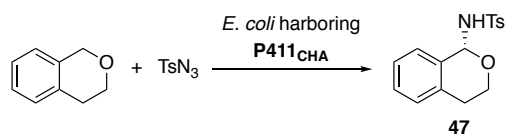
R47, Y51: Polar residues that interact with the carboxylate of the native fatty acid substrates

**Table B-6.** Intermolecular C–H amination data presented in Figure 3-5.<sup>a</sup>

Variant	Substrate	[P411]	Yield	ee	TTN
P-4	4-ethylanisole	1.8 $\mu$ M	11 $\pm$ 1%	14% ( <i>S</i> )	310
P-4 A82L	4-ethylanisole	2.4 $\mu$ M	51 $\pm$ 3%	77% ( <i>R</i> )	1000
P-4 A82L A78V	4-ethylanisole	2.7 $\mu$ M	66 $\pm$ 2%	80% ( <i>R</i> )	1200
P-4 A82L A78V F263L	4-ethylanisole	N.A. <sup>a</sup>	66 $\pm$ 2%	>99% ( <i>R</i> )	N.A. <sup>a</sup>
P411 <sub>CHA</sub>	4-ethylanisole	3.2 $\mu$ M	66 $\pm$ 3%	>99% ( <i>R</i> )	1000
P-4	4-ethyltoluene	1.8 $\mu$ M	2.0 $\pm$ 0.3%	N.A.	58
P-4 A82L	4-ethyltoluene	2.4 $\mu$ M	11 $\pm$ 1%	98% ee ( <i>R</i> )	220
P-4 A82L A78V	4-ethyltoluene	2.7 $\mu$ M	29 $\pm$ 1%	>99% ee ( <i>R</i> )	530
P-4 A82L A78V F263L	4-ethyltoluene	N.A. <sup>a</sup>	32 $\pm$ 1%	>99% ee ( <i>R</i> )	N.A. <sup>a</sup>
P411 <sub>CHA</sub>	4-ethyltoluene	3.2 $\mu$ M	34 $\pm$ 3%	>99% ee ( <i>R</i> )	530
P-4	ethylbenzene	1.8 $\mu$ M	0.50 $\pm$ 0.01%	N.A.	15
P-4 A82L	ethylbenzene	2.4 $\mu$ M	2.2 $\pm$ 0.1%	N.A.	46
P-4 A82L A78V	ethylbenzene	2.7 $\mu$ M	6.5 $\pm$ 0.4%	92% ee ( <i>R</i> )	120
P-4 A82L A78V F263L	ethylbenzene	N.A. <sup>a</sup>	6.7 $\pm$ 0.4%	>99% ee ( <i>R</i> )	N.A. <sup>a</sup>
P411 <sub>CHA</sub>	ethylbenzene	3.2 $\mu$ M	15 $\pm$ 1%	>99% ee ( <i>R</i> )	240

<sup>a</sup>Reactions were performed using *E. coli* cells expressing the P411 variant, resuspended to OD<sub>600</sub> = 30, with 5 mM alkane substrate and 5 mM tosyl azide, under anaerobic conditions. Results are the average of experiments performed with duplicate cell cultures, each used to perform duplicate chemical reactions (four reactions total). <sup>b</sup>Protein concentration could not be accurately determined; variant is poorly behaved in carbon monoxide-binding assay.



**Table B-7.** Enantioselective C–H amination of isochroman with cytochrome P411<sub>CHA</sub> followed by post-reaction racemization.<sup>a</sup>

Entry	Conditions	ee
1	1 hour reaction time	65%
2	2 hour reaction time	65%
3	4 hour reaction time	65%
4	6 hour reaction time	65%
5	20 hour reaction time	65%
6	20 hour reaction time, followed by treatment with silica	racemic

<sup>a</sup>Reactions performed as in Table 2, using whole *E. coli* cells overexpressing P411<sub>CHA</sub> at  $\text{OD}_{600} = 30$  ( $\sim 3 \mu\text{M}$  enzyme), with 2.5 mM alkane and 5 mM tosyl azide. Identical small scale reactions (400  $\mu\text{L}$ ) were set up in parallel. After the indicated time, crude reaction mixtures were extracted with 1:1 cyclohexane:ethyl acetate, concentrated to dryness, and redissolved in 50% isopropanol in cyclohexane. Silica treatment was performed by allowing the crude reaction mixture (in cyclohexane/ethyl acetate) to rest on a silica pipette column (approx. 1 hour) followed by elution with ethyl acetate. Samples were analyzed by SFC chromatography using Chiralpak AS column (35% isopropanol). Representative traces are shown below.

**Representative traces to accompany Table B-7.**

Racemic standard, *N*-tosyl-isochroman-1-amine (*rac*-**47**):

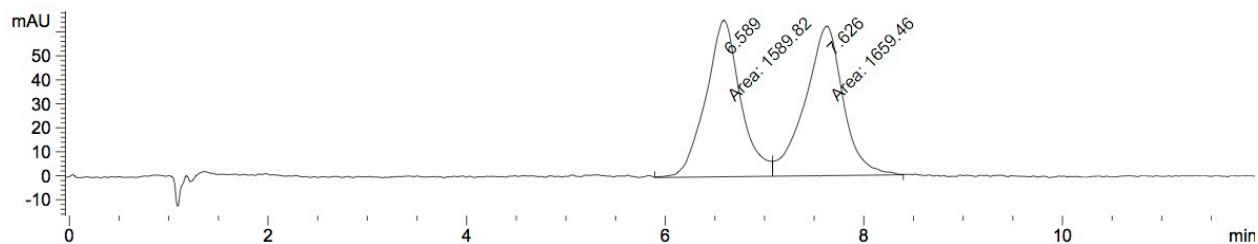


Table B-7, Entry 5, with P411<sub>CHA</sub> (65% ee):

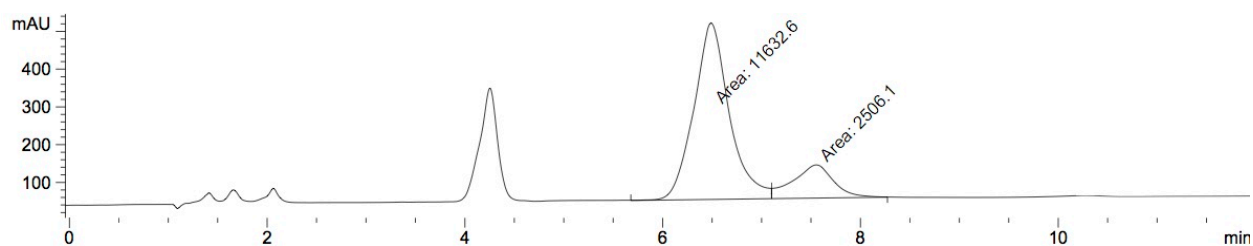
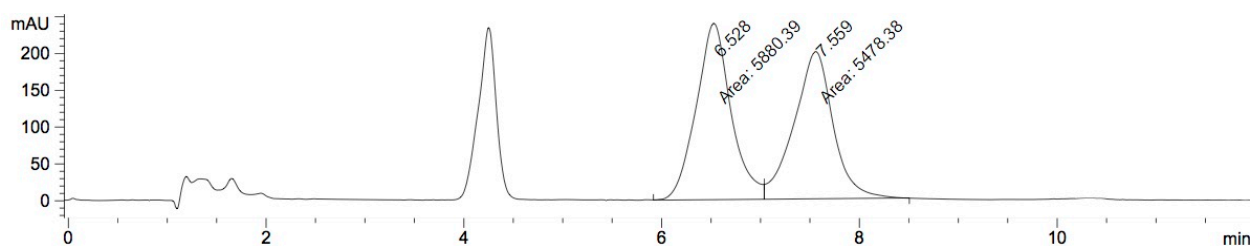
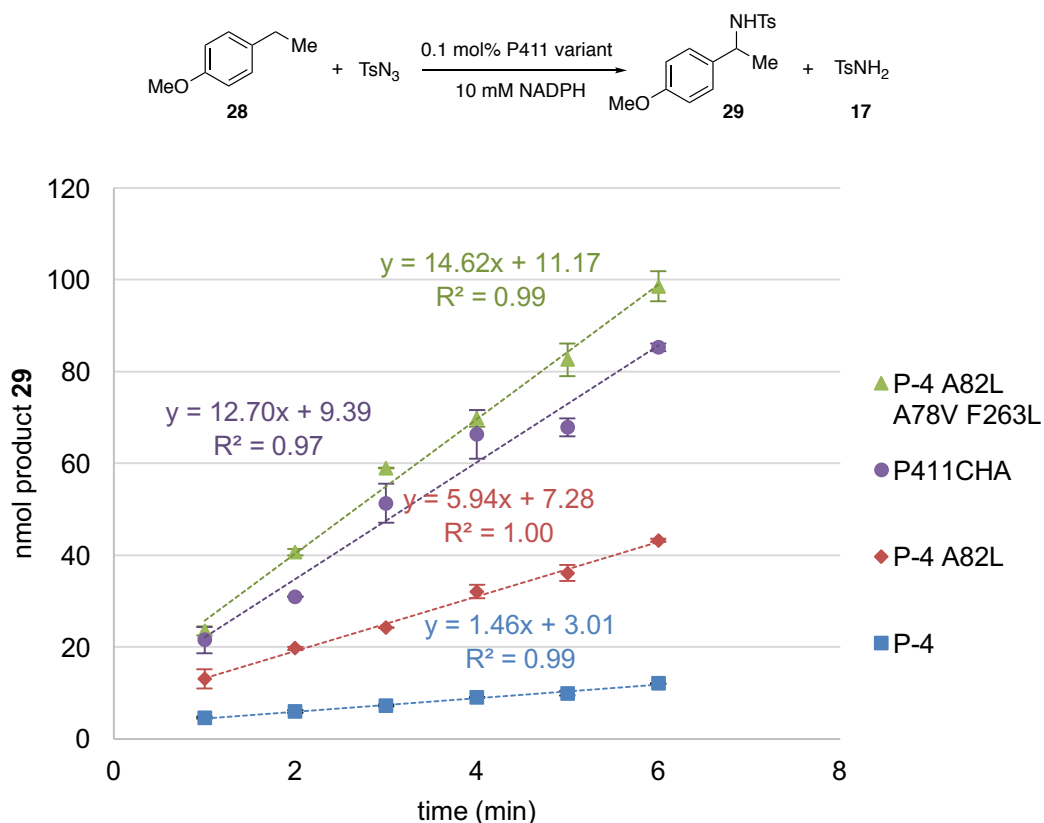


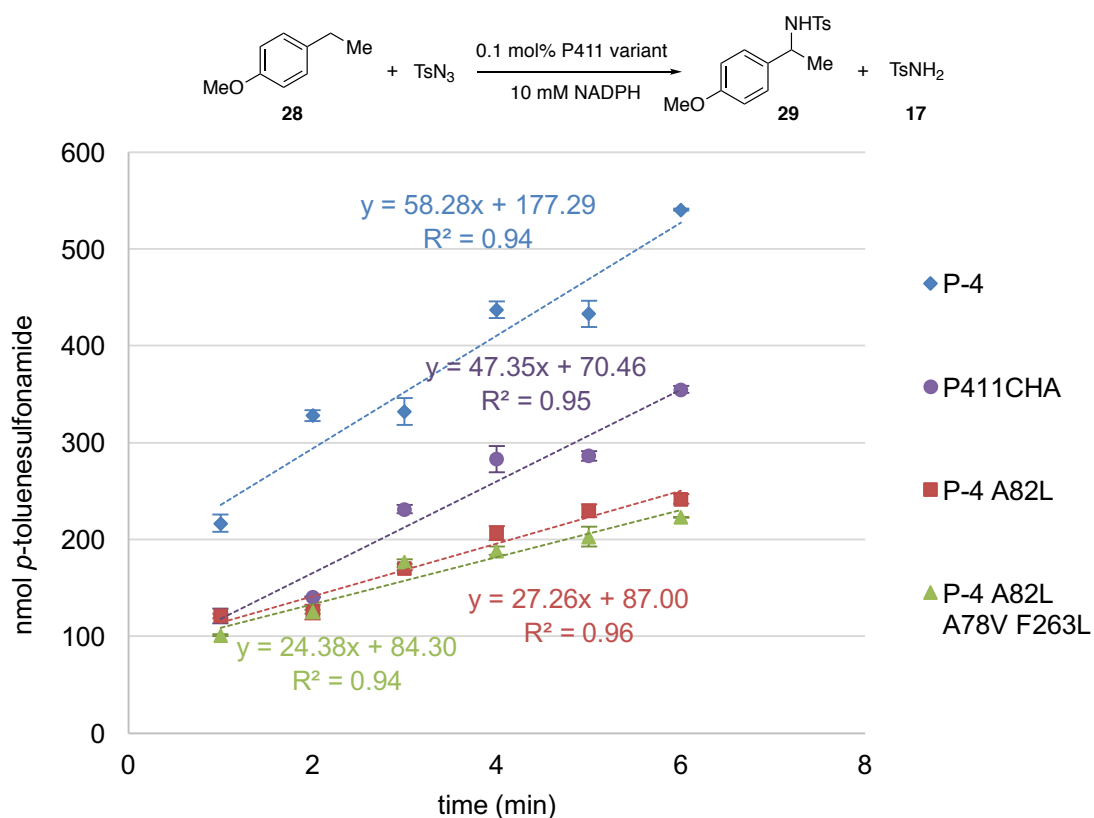
Table B-7, Entry 6, with P411<sub>CHA</sub> post silica treatment (racemic):



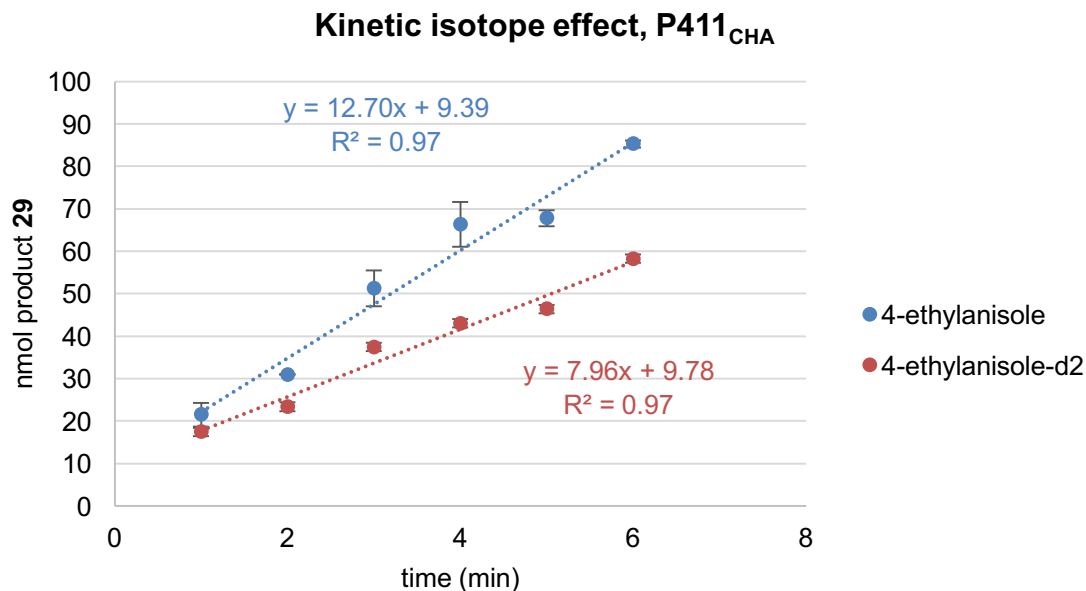
## B.2 Supporting Figures B-1 through B-3



**Figure B-1.** Initial rates of intermolecular C–H amination of 4-ethylanisole. Reactions performed with 5 mM 4-ethylanisole, 5 mM tosyl azide, 10 mM NADPH and 5  $\mu$ M purified enzyme; see Experimental Methods (Section 3.5) for details. These data were used to calculate the results presented in Table 3-2.



**Figure B-2.** Initial rates of azide reduction in the C–H amination of 4-ethylanisole. Reactions performed with 5 mM 4-ethylanisole, 5 mM tosyl azide, 10 mM NADPH and 5  $\mu$ M purified enzyme; see Experimental Methods (Section 3.5) for details. These data were obtained from the same experiments as Figure B-1 and used to calculate the results presented in Table 3-2.



**Figure B-3.** Independent rate experiments of P411<sub>CHA</sub>-catalyzed C–H amination of 4-ethylanisole or 4-ethylanisole-*d*<sub>2</sub> with tosyl azide. Reactions performed with 5 mM 4-ethylanisole or 5 mM 4-ethylanisole-*d*<sub>2</sub>, 5 mM tosyl azide, 10 mM NADPH and 5  $\mu$ M P411<sub>CHA</sub> purified enzyme; see Experimental Methods (Section 3.5) for details. These data were used to calculate a kinetic isotope effect ( $k_H/k_D$ ) of 1.6, which was presented in Figure 3-7.

### B.3 Amino Acid Sequences

#### Amino acid sequence of cytochrome P411<sub>CHA</sub>:

MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADDELGEIFKFEAPGRVTRYLSSQRLIKEACDE  
SRFDKNLSQALKFVRDFLGDLATSWTHEKNWKKAHNILLPSFSQQAMKGYHAMMVDIAVQLVQK  
WERLNADEHIEVSEDMTRLTLDTIGLCGFNYRFNSFYRDQPHFIIISMVRALDEV MNKLQRANPD  
DPAYDENKRQFQEDIKVMNDLVDKIIADRKARGEQSDDLTLQMLNGKDPETGEPLDDGNIRYQII  
TFLLAGHDGTSGLLSFALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRL  
WPTVPAFSLYAKEDTVLGGEYPLEKGDEV MLIPQLHRDKTVWGDDVEEFRPERFENPSAIPQHA  
FKPFGNGQRASIGQQFALHEATLVLGMMMLKHFD FEDHTNYELDIKETLSLKP KGFVVKAKSKKIP  
LGGIPSPSTEQSAKKVRKKAENAHNTPLLVLVYGSNMGTAE GTARDLADIAMSKGFAPQVATLDSH  
AGNLPREGAVLIVTASYN GHPPDNAKQFVDWLDQASADEVKGVRYSVFGCGDKNWATTYQKVPAF  
IDETLAAGAENIADRGEADASDDFEGTYEEWREH MWSDVAAYFNLDIENSEDNKSTLSLQFVDS  
AADMPLAKMHGAFSTNVVASKELQQPGSARSTRHLEIELPKEASYQEGDHLGVIPRNYEGIVNRV  
TARFGLDASQQIRLEAEE EKLAHLPLAKTVSVEELLQYVELQDPVTRTQLRAMAAKTVCPPHKVE  
LEALLEKQAYKEQVLAKRLT MLELLEKYPACEMKFSEFIAL LPSIRPRYYSISSSPRVDEKQASI  
TVSVVSGEAWSGYGEYKGIASNYLAELQEGDTITCFISTPQSEFTLPKDPETPLIMVPGTGVPAP  
FRGFVQARKQLKEQGQSLGEAHL YFGCRSPHEDYLYQEEL ENAQSEGIITLHTAFSRMPNQPKTY

VQHVMEQDGKKLIELLDQGAHFYICGDGSQMAPAVEATLMKSYADVHQVSEADARLWLQQLEEKGRYAKDVWAG

### Sequence of heme domain of P-4 A82L A78V F263L used for crystallography:

The construct employed for crystallization is the heme domain of variant P-4 A82L A78V F263L (residues 1 to 463 of the holoprotein, followed by an *Xho*I site and a C-terminal His-tag).

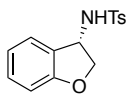
MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADDELGEIFKFEAPGRVTRYLSSQRLIKEACDESRFDKNLSQALKFVRDFLDGLATSWTHEKNWKKAHNILLPSFSQQAMKGYHAMMVDIAVQLVQKWERLNADEHIEVSEDMTRLTLDTIGLCGFNYRFNSFYRDQPHFPIISMVRALDEVMNKLQRANPD DPAYDENKRQFQEDIKVMNDLVDKIIADRKARGEQSDDLTLQMLNGKDPETGEPLDDGNIRYQII TFLLAGHEGTSGLLSFALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRL WPTVPAFSLYAKEDTVLGGEYPLEKGDEVMLIPQLHRDKTVWGDDVEEFRPERFENPSAIPQHA FKPFNGQQRASIGQQFALHEATLVLGMMMLKHFD FEDHTNYELDIKETLSLKP KGFVVKAKSKKIP LGGIPSPSTLEHHHHHH

## B.4 Synthesis and Characterization of Substrates and Reference Compounds

### B.4.1 Substrates and Reference Compounds for Sections 3.3.1 to 3.3.5

4-Ethylanisole- $d_2$  (**28- $d_2$** ) was prepared according the procedure of Kurita *et al.*<sup>6</sup> to 97% deuterium incorporation at the benzylic position, as determined by  $^1\text{H}$  NMR analysis.

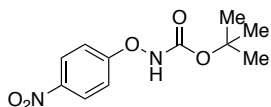
Reference compounds corresponding to enzymatic reaction products were prepared by tosylation of the corresponding benzylic amine or by racemic C–H amination.<sup>7</sup> The following are known compounds and their spectral data are in agreement with reported values<sup>7–12</sup>: *N*-tosyl-1-(*p*-methoxyphenyl)ethylamine (**29**), *N*-tosyl-1-(*p*-tolyl)ethylamine (**32**), *N*-tosyl-1-(*m*-tolyl)ethylamine (**33**), *N*-tosyl-1-(*o*-tolyl)ethylamine (**34**), *N*-tosyl-1-phenylethylamine (**35**), *N*-tosyl-1-(*p*-bromophenyl)ethylamine (**36**), *N*-Tosyl-1-(*p*-chlorophenyl)ethylamine (**37**), *N*-tosyl-1-(*p*-fluorophenyl)ethylamine (**38**), *N*-tosyl-1-aminoindane (**39**), *N*-tosyl-1-aminotetralin (**40**), *N*-tosyl-1-(2-naphthyl)ethylamine (**42**), *N*-tosyl-1-(1-naphthyl)ethylamine (**43**), *N*-tosyl-1-(*p*-methoxyphenyl)propylamine (**44**), *N*-tosyl-4-methoxybenzylamine (**45**), *N*-tosyl-1,3-dihydroisobenzofuran-1-amine (**46**), *N*-tosyl-isochroman-1-amine (**47**).  $^1\text{H}$  NMR data for these compounds are provided in the Supplementary Information of the published paper.

***N*-Tosyl-3-amino-2,3-dihydrobenzofuran (41)**

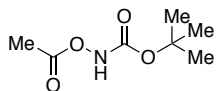
**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.81 (d, *J* = 8.3 Hz, 2H), 7.37 (d, *J* = 7.9 Hz, 2H), 7.23–7.17 (m, 1H), 6.94–6.88 (m, 1H), 6.86–6.78 (m, 2H), 5.05 (ddd, *J* = 8.1, 8.1, 4.1 Hz, 1H), 4.77 (d, *J* = 8.5 Hz, 1H), 4.49 (dd, *J* = 10.3, 7.9 Hz, 1H), 4.28 (dd, *J* = 10.3, 4.1 Hz, 1H), 2.48 (s, 3H); **<sup>13</sup>C NMR** (100 MHz, CDCl<sub>3</sub>) δ 160.0, 144.1, 137.8, 130.9, 130.2, 127.3, 125.3, 125.0, 121.4, 110.7, 77.4, 55.6, 21.8; **IR** (film) 3284, 2921, 1598, 1482, 1330, 1155, 1082, 970, 750, 667 cm<sup>-1</sup>; **HRMS** (FAB<sup>+</sup>) exact mass calculated for C<sub>15</sub>H<sub>16</sub>NO<sub>3</sub>S<sup>+</sup> requires *m/z* 290.0851, found 290.0855.

**B.4.1 Substrates and Reference Compounds for Section 3.3.6**

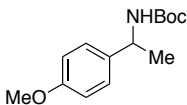
Substrates and reference compounds associated with Section 3.3.6 were prepared according to literature procedures.

***tert*-butyl (4-nitrophenoxy)carbamate (48)**

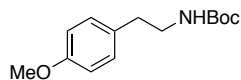
This compound was prepared following the method of T. Sheradsky *et al.*<sup>13</sup> **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 8.22 (d, *J* = 9.3 Hz, 2H), 7.22 (d, *J* = 9.3 Hz, 2H), 1.51 (s, 9H).

***tert*-butyl acetoxycarbamate (49).**

This compound was prepared following the method of P. Patel *et al.*<sup>14</sup>; spectral data are in agreement with literature reported values.<sup>15</sup> **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.90 (s, 1H), 2.20 (s, 3H), 1.49 (s, 9H).

***tert*-butyl (1-(4-methoxyphenyl)ethyl)carbamate (50)**

This compound was prepared from 1-(4-methoxyphenyl)ethan-1-amine and di-*tert*-butyl dicarbonate following the method of F. Jahani *et al.*<sup>16</sup> Spectral data are in agreement with literature reported values.<sup>17</sup> **<sup>1</sup>H NMR** (300 MHz, CDCl<sub>3</sub>) δ 7.22 (d, *J* = 8.4 Hz, 2H), 6.87 (d, *J* = 8.6 Hz, 2H), 4.73 (br s, 2H), 3.79 (s, 3H), 1.44 (d, *J* = 6.8 Hz, 3H), 1.42 (s, 9H).

***tert*-butyl (4-methoxyphenethyl)carbamate (50')**

This compound was prepared from 2-(4-methoxyphenyl)ethan-1-amine and di-*tert*-butyl dicarbonate; spectral data are in agreement with literature reported values.<sup>16</sup> <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.11 (d, *J* = 8.5 Hz, 2H), 6.85 (d, *J* = 8.6 Hz, 2H), 4.52 (br s, 1H), 3.79 (s, 3H), 3.34 (q, *J* = 6.7 Hz, 2H), 2.73 (t, *J* = 7.0 Hz, 2H), 1.43 (s, 9H).

**B.5 HPLC Calibration**

Calibration curves with an internal standard were created for quantitative HPLC analysis of reaction products; these curves plot the concentration of analyte (mM, y-axis) against the HPLC ratio (peak area ratio of analyte over internal standard, x-axis). Internal standards were 1,3,5-trimethoxybenzene (1.25 mM) or methyl phenylacetate (1.0 mM) and reactions were monitored at 230 nm or 210 nm. Calibration curves can be found in the Supplementary Information of the published paper. The identity of the products was additionally confirmed by HPLC co-injections of reaction mixtures with chemically synthesized authentic products, or by NMR analysis of products isolated from reactions performed on preparative scale.

**B.6 Preparative Scale Reactions**

These experiments were conducted by Dr. Chris Prier. *E. coli* BL21 *E. cloni* cells transformed with the plasmid encoding P411<sub>CHA</sub> were grown overnight in 25 mL LB<sub>amp</sub> (37 °C, 250 rpm). Hyperbroth medium (470 mL, 0.1 mg/mL ampicillin) in a 1-L flask was inoculated with 19 mL of the preculture and incubated at 37 °C and 230 rpm for 2.5 h (to OD<sub>600</sub> ca. 1.8). Cultures were then cooled on ice (30 min) and induced with 0.5 mM IPTG and 1.0 mM 5-aminolevulinic acid (final concentrations). Expression was conducted at room temperature (23 °C) at 130 rpm for 16–18 h (to OD<sub>600</sub> ca. 7.0). Cultures were then centrifuged (2,600 x g, 10 min, 4 °C) and the pellets resuspended to OD<sub>600</sub> = 30 in M9-N. Aliquots of the cell suspension (4 mL) were used to determine the P411 expression level after lysis by sonication. The cells (80 mL) were then combined with glucose (10 mL, 250 mM in M9-N) in a 250-mL Erlenmeyer flask and degassed by sparging with argon for at



least 40 minutes. The reaction flask was then transferred into an anaerobic chamber. To the flask were added the oxygen depletion system (5 mL, 14,000 U/mL catalase and 1,000 U/mL glucose oxidase in 0.1 M KPi, pH 8.0) followed by alkane (2.5 mL, 100 mM in DMSO) and tosyl azide (2.5 mL, 200 mM in DMSO). Final concentrations were typically 2.5 mM alkane, 5.0 mM tosyl azide, and 25 mM glucose; final reaction volume was 100 mL. The flask was sealed with parafilm, removed from the anaerobic chamber, and shaken at room temperature and 130 rpm for 20 h. The reaction was quenched by adding acetonitrile (50 mL) and then centrifuged (4,000 x g, 10 min). The supernatant was concentrated and extracted with EtOAc (3 x 25 mL). The organic layers were washed with brine (20 mL), dried over MgSO<sub>4</sub>, filtered, concentrated, and purified by chromatography.

***N*-Tosyl-1-(*p*-methoxyphenyl)ethylamine (29).** The reaction was performed under two different conditions. To optimize yield, the reaction was performed on 0.25-mmol scale with 2.5 mM 4-ethylanisole, 5.0 mM tosyl azide, and 3.18  $\mu$ M P411<sub>CHA</sub>. The product was purified by silica gel chromatography (5 to 30% EtOAc/hexanes). Isolated 59.5 mg (78% yield, 610 TTN, >99% ee). To optimize turnovers, the reaction was performed on 1.0-mmol scale with 10 mM 4-ethylanisole, 10 mM tosyl azide, and 2.80  $\mu$ M P411<sub>CHA</sub>. The product was purified by silica gel chromatography (20 to 30% EtOAc/hexanes). Isolated 103.3 mg (34% yield, 1,200 TTN, >99% ee).

***N*-Tosyl-3-amino-2,3-dihydrobenzofuran (41).** The reaction was performed on 0.25-mmol scale with 2.5 mM 2,3-dihydrobenzofuran, 5.0 mM tosyl azide, and 2.98  $\mu$ M P411<sub>CHA</sub>. The product was purified by C18 chromatography (5 to 100% MeCN/water). Isolated 15.1 mg (21% yield, 180 TTN, 92% ee).

***N*-Tosyl-1,3-dihydroisobenzofuran-1-amine (46).** The reaction was performed on 0.25-mmol scale with 2.5 mM *o*-xylylene oxide, 5.0 mM tosyl azide, and 2.86  $\mu$ M P411<sub>CHA</sub>. The product was purified by preparative HPLC (50 to 100% MeCN/water). Isolated 50.7 mg (70% yield, 610 TTN).

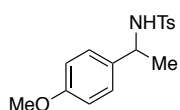
***N*-Tosyl-isochroman-1-amine (47).** The reaction was performed on 0.25-mmol scale with 2.5 mM isochroman, 5.0 mM tosyl azide, and 2.73  $\mu$ M P411<sub>CHA</sub>. The product was purified by silica gel chromatography (5 to 30% EtOAc/hexanes). Isolated 60.6 mg (80% yield, 730 TTN).

## B.7 Determination of Enantioselectivity

For the determination of enantioselectivity, reaction mixtures were extracted with cyclohexane, or purified compounds were dissolved in 20% isopropanol in hexanes, and samples were analyzed by chiral SFC.

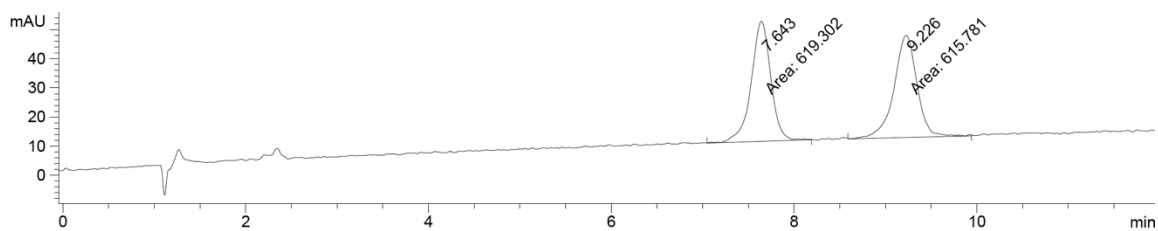
### B.7.1 Assignment of Absolute Stereochemistry

Assignment of absolute stereochemistry was performed by Dr. Chris Prier. Absolute stereochemistry was assigned by chemical synthesis of (*R*)-*N*-tosyl-1-(*p*-methoxyphenyl)ethylamine ((*R*)-**29**), (*S*)-*N*-tosyl-1-(*p*-tolyl)ethylamine ((*S*)-**32**), and (*R*)-*N*-tosyl-1-(1-naphthyl)ethylamine ((*R*)-**43**). Comparison to the enzymatic reactions reveals the enzymatic products to be the *R* enantiomer. Other benzylic amine products were assigned by analogy. Traces for the assignment using compound **29** are shown; additional traces for the other compounds can be found in the Supporting Information of the published paper.

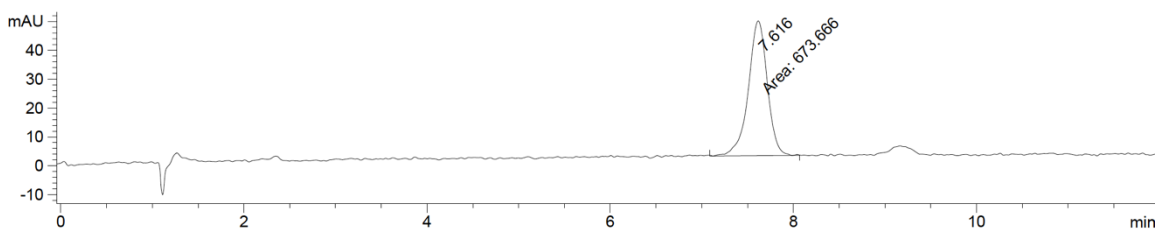


***N*-Tosyl-1-(*p*-methoxyphenyl)ethylamine (29).** SFC (Chiralpak AS column): 25% isopropanol in supercritical CO<sub>2</sub>, 2.5 mL/min, 40 °C.

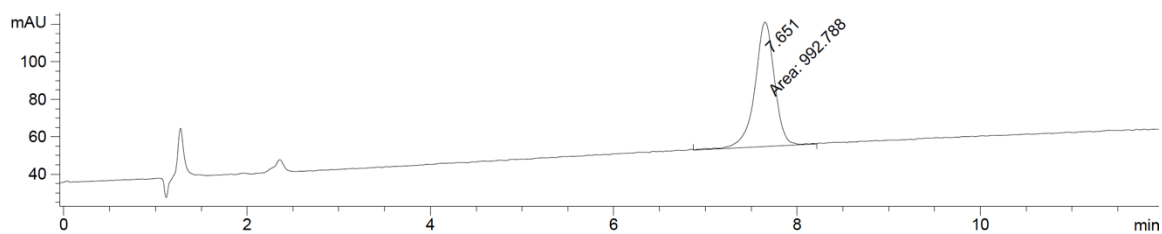
*rac*-29



(*R*)-29, chemically prepared:

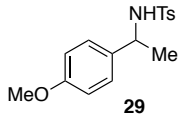
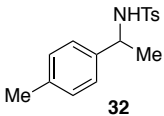
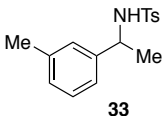
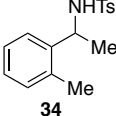
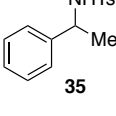
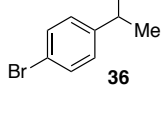
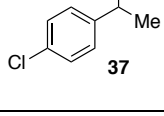
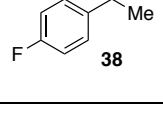


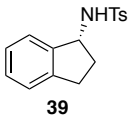
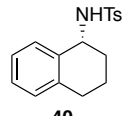
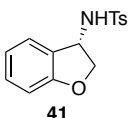
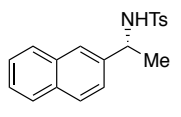
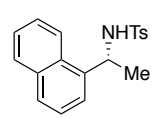
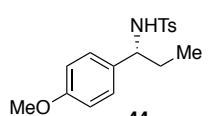
Enzymatic reaction, with P411<sub>CHA</sub> (>99% ee):



### B.7.2 Analytical Conditions for Measurement of Enantiomeric Excess

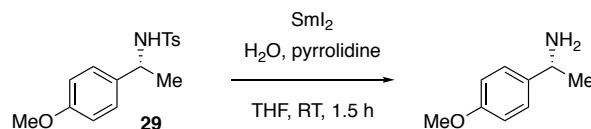
Representative traces for P411<sub>CHA</sub> catalyzed production of compounds **29**, **32–44** are included in the Supporting Information of the published paper.

Compound	Analytical conditions for separating enantiomers
 <b>29</b>	SFC (Chiralpak AS column): 25% isopropanol in supercritical CO <sub>2</sub> , 2.5 mL/min, 40 °C t <sub>R</sub> : 7.6 min, 9.2 min
 <b>32</b>	SFC (Chiralpak AS column): 25% isopropanol in supercritical CO <sub>2</sub> , 2.5 mL/min, 40 °C t <sub>R</sub> : 6.3 min, 7.8 min
 <b>33</b>	SFC (Chiralpak AS column): 25% isopropanol in supercritical CO <sub>2</sub> , 2.5 mL/min, 40 °C t <sub>R</sub> : 5.5 min, 8.0 min
 <b>34</b>	SFC (Chiralpak AS column): 30% isopropanol in supercritical CO <sub>2</sub> , 2.5 mL/min, 40 °C t <sub>R</sub> : 5.0 min, 8.2 min
 <b>35</b>	SFC (Chiralpak AS column): 25% isopropanol in supercritical CO <sub>2</sub> , 2.5 mL/min, 40 °C t <sub>R</sub> : 6.2 min, 7.7 min
 <b>36</b>	SFC (Chiralpak AS column): 30% isopropanol in supercritical CO <sub>2</sub> , 2.5 mL/min, 40 °C t <sub>R</sub> : 5.1 min, 6.7 min
 <b>37</b>	SFC (Chiralpak AS column): 25% isopropanol in supercritical CO <sub>2</sub> , 2.5 mL/min, 40 °C t <sub>R</sub> : 5.9 min, 8.1 min
 <b>38</b>	SFC (Chiralpak AS column): 25% isopropanol in supercritical CO <sub>2</sub> , 2.5 mL/min, 40 °C t <sub>R</sub> : 4.6 min, 6.3 min

 <b>39</b>	SFC (Chiralpak AS column): 35% isopropanol in supercritical CO <sub>2</sub> , 2.5 mL/min, 40 °C t <sub>R</sub> : 6.1 min, 8.4 min
 <b>40</b>	SFC (Chiralpak AS column): 40% isopropanol in supercritical CO <sub>2</sub> , 2.5 mL/min, 40 °C t <sub>R</sub> : 5.2 min, 7.3 min
 <b>41</b>	SFC (Chiralpak AS column): 30% isopropanol in supercritical CO <sub>2</sub> , 2.5 mL/min, 40 °C t <sub>R</sub> : 5.6 min, 8.5 min
 <b>42</b>	SFC (Chiralpak AS column): 40% isopropanol in supercritical CO <sub>2</sub> , 2.5 mL/min, 40 °C t <sub>R</sub> : 4.3 min, 5.9 min
 <b>43</b>	SFC (Chiralpak AS column): 40% isopropanol in supercritical CO <sub>2</sub> , 2.5 mL/min, 40 °C t <sub>R</sub> : 4.5 min, 8.0 min
 <b>44</b>	SFC (Chiralpak AS column): 25% isopropanol in supercritical CO <sub>2</sub> , 2.5 mL/min, 40 °C t <sub>R</sub> : 6.3 min, 9.3 min

## B.8 Tosyl Group Removal

Tosyl group removal was performed by the method of Ankner and Hilmersson.<sup>18</sup>

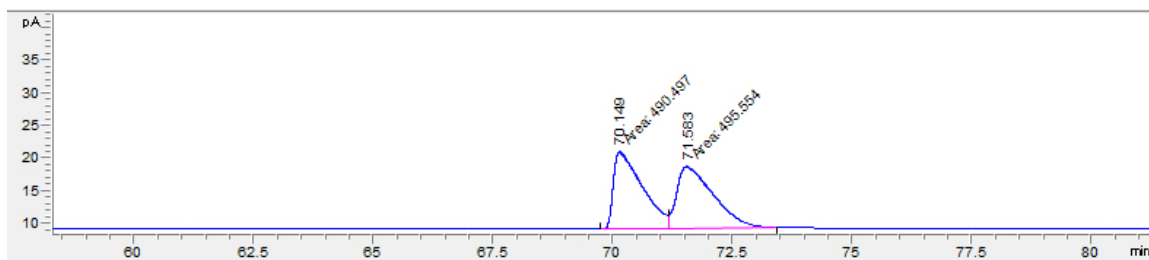


Enzymatically produced tosylamide **29** (30.5 mg, 0.1 mmol, >99% ee) in dry THF (1 mL) was added to a 0.1 M solution of SmI<sub>2</sub> in THF (20 mL, 2.0 mmol) at room temperature under argon. Water (108 µL, 6.0 mmol, degassed for 10 min with argon) and pyrrolidine (334 µL, 4.0 mmol, degassed for 10 min with argon) were then added. The solution was

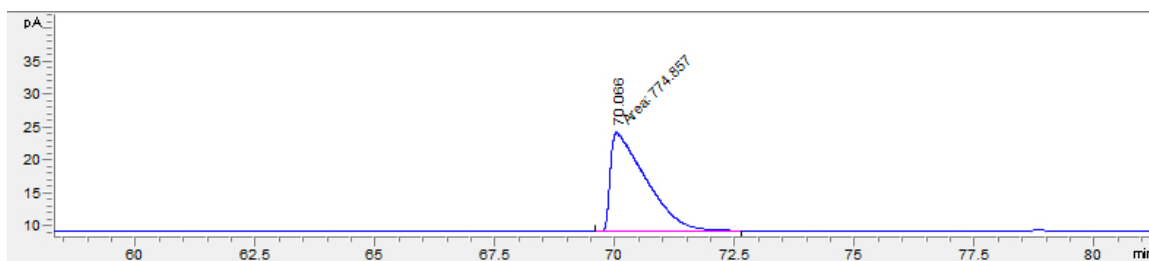
allowed to stir at room temperature under argon for 1.5 hours. The resulting reaction mixture was diluted with diethyl ether (20 mL) and quenched with a solution of potassium sodium tartrate and potassium carbonate (20 mL, 10% w/v each). The aqueous phase was extracted with three portions of diethyl ether ( $3 \times 20$  mL). The combined organics were pooled and evaporated to yield the crude amine. The crude amine was taken up in aqueous 10% HCl solution (5 mL) and washed with diethyl ether ( $4 \times 3$  mL). The aqueous layer was basified with 10% NaOH solution (5 mL) to pH 14 and extracted with dichloromethane ( $3 \times 10$  mL). The combined organic layers were dried over  $\text{Na}_2\text{SO}_4$ . Concentration gave (*R*)-1-(4-methoxyphenyl)ethanamine (9.2 mg, 61% yield) as an oil with spectral data in agreement with reported values.<sup>19</sup> The ee was determined to be >99% using an Agilent 7820A GC equipped with a Cyclosil-B chiral column (30 m  $\times$  0.320 mm, 0.25  $\mu\text{m}$  film); method: 90–98 °C at 0.1 °C/min, 98–240 °C at 15 °C/min, 240 °C hold 5 min.

**(*R*)-1-(4-methoxyphenyl)ethanamine:**  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.26 (d,  $J = 8.5$  Hz, 2H), 6.87 (d,  $J = 8.7$  Hz, 2H), 4.08 (q,  $J = 6.6$  Hz, 1H), 3.80 (s, 3H), 1.59 (br s, 2H), 1.36 (d,  $J = 6.6$  Hz, 3H).

**Representative traces.** Racemic standard, 1-(4-methoxyphenyl)ethanamine:



(*R*)-1-(4-methoxyphenyl)ethanamine from tosyl group removal of enzymatically produced **29**:



**B.9 References and Notes**

1. Hyster, T. K.; Farwell, C. C.; Buller, A. R.; McIntosh, J. A.; Arnold, F. H. Enzyme-Controlled Nitrogen-Atom Transfer Enables Regiodivergent C–H Amination. *J. Am. Chem. Soc.* **2014**, *136*, 15505–15508.
2. Farwell, C. C.; Zhang, R. K.; McIntosh, J. A.; Hyster, T. K.; Arnold, F. H. Enantioselective enzyme-catalyzed aziridination enabled by active-site evolution of a cytochrome P450. *ACS Cent. Sci.* **2015**, *1*, 89–93.
3. Bordeaux, M.; Singh, R.; Fasan, R. Intramolecular C(sp<sup>3</sup>)–H amination of arylsulfonyl azides with engineered and artificial myoglobin-based catalysts. *Bioorg. Med. Chem.* **2014**, *22*, 5697–5704.
4. Kan, S. B. J.; Lewis, R. D.; Chen, K.; Arnold, F. H. Directed evolution of cytochrome *c* for carbon–silicon bond formation: Bringing silicon to life. *Science* **2016**, *354*, 1048–1051.
5. Prier, C. K.; Hyster, T. K.; Farwell, C. C.; Huang, A.; Arnold, F. H. Asymmetric enzymatic synthesis of allylic amines: A sigmatropic rearrangement strategy. *Angew. Chem. Int. Ed.* **2016**, *55*, 4711–4715.
6. Kurita, T.; Hattori, K.; Seki, S.; Mizumoto, T.; Aoki, F.; Yamada, Y.; Ikawa, K.; Maegawa, T.; Monguchi, Y.; Sajiki, H. Efficient and convenient heterogeneous palladium-catalyzed regioselective deuteration at the benzylic position. *Chem. Eur. J.* **2008**, *14*, 664–673.
7. Takeda, Y.; Hayakawa, J.; Yano, K.; Minakata, S. Transition-metal-free benzylic C–H bond intermolecular amination utilizing chloramine-T and I<sub>2</sub>. *Chem. Lett.* **2012**, *41*, 1672–1674.
8. Nishimura, T.; Yasuhara, Y.; Hayashi, T. Asymmetric addition of dimethylzinc to *N*-tosylarylimines catalyzed by a rhodium-diene complex toward the synthesis of chiral 1-arylethylamines. *Org. Lett.* **2006**, *8*, 979–981.
9. Taylor, J. G.; Whittall, N.; Hii, K. K. Copper-Catalyzed Intermolecular Hydroamination of Alkenes. *Org. Lett.* **2006**, *8*, 3561–3564.
10. Albone, D. P.; Challenger, S.; Derrick, A. M.; Fillery, S. M.; Irwin, J. L.; Parsons, C. M.; Takada, H.; Taylor, P. C.; Wilson, D. J. Amination of ethers using Chloramine-T hydrate and a copper(I) catalyst. *Org. Biomol. Chem.* **2005**, *3*, 107–111.

- 
11. Harden, J. D.; Ruppel, J. V.; Gao, G.-Y.; Zhang, X. P. Cobalt-catalyzed intermolecular C–H amination with Bromamine-T as nitrene source. *Chem. Commun.* **2007**, 4644–4646.
  12. Molander, G. A.; Fleury-Brégeot, N.; Hiebel, M.-A. Synthesis and cross-coupling of sulfonamidomethyltrifluoroborates. *Org. Lett.* **2011**, *13*, 1694–1697.
  13. Sheradsky, T.; Salemnick, G.; Nir, Z. Introduction of the aminooxy group on to nitroaromatic and heterocyclic rings : Synthesis and properties of O-(nitroaryl)hydroxylamines. *Tetrahedron* **1972**, *28*, 3833–3843.
  14. Pitambar, P.; Chang, S. Cobalt(III)-catalyzed C–H amidation of arenes using acetoxycarbamates as convenient amino sources under mild conditions. *ACS Catal.* **2015**, *5*, 853–858.
  15. Pitambar, P.; Chang, S. *N*-Substituted hydroxylamines as synthetically versatile amino sources in the Iridium-catalyzed mild C–H amidation reaction. *Org. Lett.* **2014**, *16*, 3328–3331.
  16. Jahani, F.; Tajbakhsh, M.; Golchoubian, H.; Khaksar, S. Guanidine hydrochloride as an organocatalyst for *N*-Boc protection of amino groups. *Tetrahedron Lett.* **2011**, *52*, 1260–1264.
  17. Ohshima, T.; Nakahara, Y.; Ipposhi, J.; Miyamoto, Y.; Mashima, K. Direct substitution of the hydroxyl group with highly functionalized nitrogen nucleophiles catalyzed by Au(III). *Chem. Commun.* **2011**, *47*, 8322–8324.
  18. Ankner, T.; Hilmersson, G. Instantaneous deprotection of tosylamides and esters with SmI<sub>2</sub>/amine/water. *Org. Lett.* **2009**, *11*, 503–506.
  19. Guijarro, D.; Pablo, O.; Yus, M. Asymmetric synthesis of chiral primary amines by transfer hydrogenation of *N*-(*tert*-butanesulfinyl)ketimines. *J. Org. Chem.* **2010**, *75*, 5265–5270.



## CARBON–CARBON BOND FORMATION BY HEME PROTEIN CATALYZED C–H FUNCTIONALIZATION

Material for this chapter appears in **Zhang, R. K.**; Chen, K.; Huang, X.; Wohlschlager, L.; Renata, H.; Arnold, F. H. “Enzymatic assembly of carbon–carbon bonds via iron-catalysed  $sp^3$  C–H functionalisation,” *Nature* **2019**, 565, 67–72. **DOI:** 10.1038/s41586-018-0808-5. This work was performed in collaboration with all authors. Reprinted with permission from Nature Publishing Group.

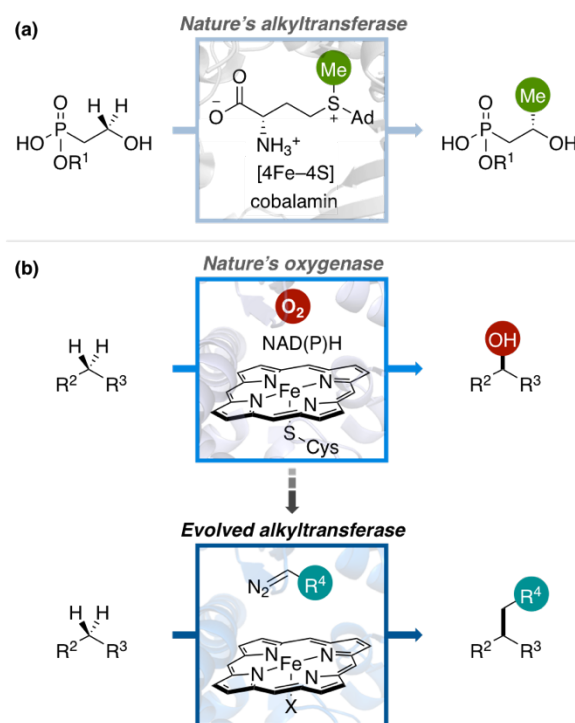
### 4.1 Abstract

Though abundant in organic molecules, carbon–hydrogen (C–H) bonds are typically considered unreactive and unavailable for chemical manipulation. Recent advances in C–H functionalization technology have begun to transform this logic, while emphasizing the challenge and importance of selective alkylation at an  $sp^3$  carbon.<sup>1,2</sup> Here we describe the first iron-based catalysts for enantioselective intermolecular alkylation of  $sp^3$  C–H bonds through carbene C–H insertion. The catalysts, derived from a cytochrome P450 enzyme whose native cysteine axial ligand has been substituted for serine (“cytochrome P411”), are fully genetically encoded and produced in bacteria, where they can be tuned by directed evolution for activity and selectivity. The laboratory-evolved enzymes functionalize diverse substrates containing benzylic, allylic, or  $\alpha$ -amino C–H bonds with high turnover and exquisite selectivity. Furthermore, these highly efficient enzymes have enabled the development of concise chemoenzymatic routes to several natural products. The demonstration that these enzymes mediate  $sp^3$  C–H alkylation using their native iron-heme cofactor unlocks a vast natural heme protein diversity for this abiological transformation and will facilitate the development of new enzymatic C–H functionalization reactions for applications in chemistry and synthetic biology.

## 4.2 Introduction

Biological systems use a limited set of chemical strategies to form carbon–carbon (C–C) bonds during construction of organic molecules.<sup>3</sup> Whereas many of these approaches rely on the manipulation of functional groups, certain enzymes, including members of the radical *S*-adenosylmethionine (SAM) family, can perform alkylation of  $sp^3$  C–H bonds. This has been an especially versatile strategy for structural diversification, as seen by its essential role in the biosynthesis of structurally varied natural products and cofactors.<sup>4–6</sup> Known biological machineries for this transformation, however, are limited to enzymes that transfer a methyl group<sup>5,6</sup> or conjugate an activated radical acceptor substrate<sup>4,7</sup> to specific molecules, with methylation as a common mode for  $sp^3$  C–alkyl installation by radical SAM enzymes (Figure 4-1a).

We sought to introduce a new enzymatic strategy for the alkylation of  $sp^3$  C–H bonds. For our design, we drew inspiration from the most widely used biological C–H functionalization transformation, C–H oxygenation. Enzymes such as the cytochromes P450 accomplish C–H oxygenation using a heme cofactor; their activities rely on activation of molecular oxygen for the controlled generation of a high-energy iron-oxo intermediate capable of selective insertion into a substrate C–H bond.<sup>8</sup> Analogously, we anticipated that the combination of a heme protein and a diazo compound would generate a protein-enclosed iron-carbene species and that this carbene could participate in a selective C–H insertion reaction with a second substrate (Figure 4-1b). While it has been shown that heme proteins are capable of performing carbene transfer processes such as cyclopropanation and heteroatom–hydrogen bond insertions<sup>9–11</sup>, their functionalization of  $sp^3$  C–H bonds remained elusive. Of note, the alkylation of  $sp^2$ -hybridized C–H bonds of unprotected indole substrates has recently been achieved using variants of myoglobin. Chemoselective for C3 functionalization, the alkylation occurs through electrophilic aromatic substitution rather than a C–H insertion mechanism which is expected for  $sp^3$  C–H functionalization.<sup>13</sup>



**Figure 4-1.** Enzymatic C–H functionalization systems. **(a)** Methylation catalyzed by cobalamin-dependent radical SAM enzymes, as illustrated by Fom3 in fosfomycin biosynthesis.<sup>6</sup> **(b)** Oxygenation catalyzed by cytochrome P450 monooxygenase (top) and envisioned alkylation reaction achieved under heme protein catalysis (bottom). Structural illustrations are adapted from Protein Data Bank (PDB) ID code 5UL4 (radical SAM enzyme) and PDB 2IJ2 (cytochrome P450<sub>BM3</sub>). Ad, adenosyl; Cys, cysteine; R, organic group; X, amino acid.

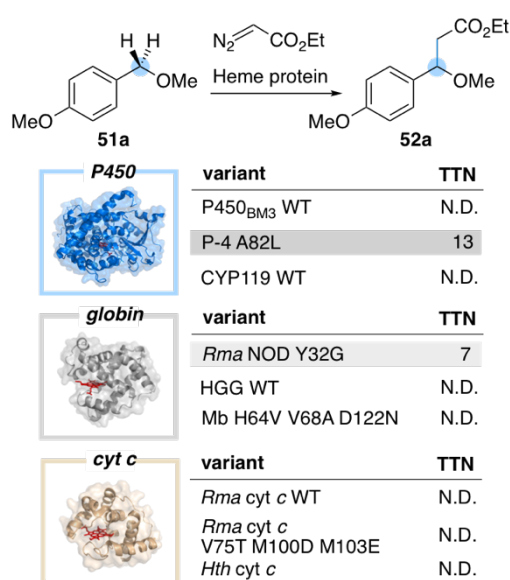
Metal-carbene  $sp^3$  C–H insertion in small-molecule catalysis, especially intermolecular and stereoselective versions of this reaction, typically relies on transition metal complexes based on rhodium,<sup>14</sup> iridium,<sup>15</sup> and others.<sup>17–19</sup> Artificial metalloproteins for carbene C–H insertion have been created by introducing an iridium-porphyrin into variants of apo heme proteins.<sup>20</sup> Though rare, there are a few examples of iron-carbene  $sp^3$  C–H insertion. The iron-catalyzed examples employ elevated temperatures (e.g. 80 °C),<sup>21</sup> are stoichiometric,<sup>22</sup> or are restricted to intramolecular reactions,<sup>23</sup> indicating a high activation energy barrier for C–H insertion with an iron-carbene. However, because the protein framework of an enzyme can impart significant rate enhancements to reactions<sup>24</sup> and even confer activity to an otherwise unreactive cofactor,<sup>25</sup> we surmised that directed

evolution could reconfigure a heme protein to overcome the barrier for the iron-carbene C–H insertion reaction and acquire this new function (Figure 4-1b).

## 4.3 Results and Discussion

### 4.3.1 Reaction discovery and directed evolution

In initial studies, we tested a panel of seventy-eight heme proteins which included variants of cytochromes P450, cytochromes *c*, and globin homologs. The heme proteins in whole *Escherichia coli* (*E. coli*) cells were combined with *p*-methoxybenzyl methyl ether (**51**) and ethyl diazoacetate at room temperature under anaerobic conditions; the resulting reactions were analysed for formation of C–H alkylation product **52a** (see Section C.3 in Appendix C for the complete list of tested heme proteins). We found heme proteins from two superfamilies that showed low levels of this promiscuous activity, establishing the possibility of creating C–H alkylation enzymes with very different protein architectures (Figure 4-2). An engineered variant of cytochrome P450<sub>BM3</sub> from *Bacillus megaterium* with an axial cysteine-to-serine mutation (cytochrome “P411”), P-4 A82L (ref. 25), provided **52a** with 13 total turnovers (TTN). In Chapter 3, P-4 A82L was found to be an effective variant for the intermolecular C–H amination of ethylbenzenes with tosyl azide. In addition, nitric oxide dioxygenase from *Rhodothermus marinus* containing the Y32G mutation (*Rma* NOD Y32G) catalyzed the reaction with 7 TTN. A second alkane substrate, 4-ethylanisole, was also accepted by the nascent C–H alkylation enzymes, albeit with lower turnover numbers (Table C-2 in Appendix C). The heme cofactor alone (iron protoporphyrin IX) or in the presence of bovine serum albumin were inactive (Table C-1 and Table C-2 in Appendix C).

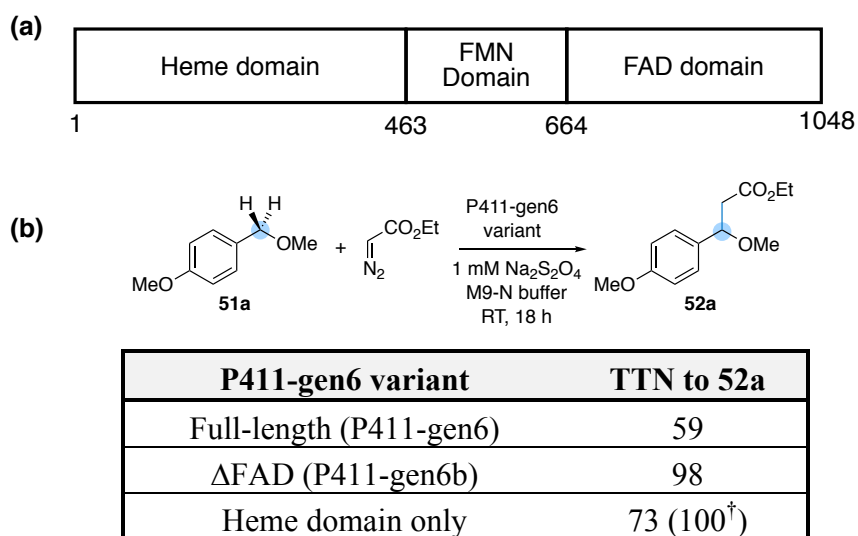


**Figure 4-2.** Select subset of heme proteins tested for promiscuous  $sp^3$  C–H alkylation activity. Structural illustrations are of representative superfamily members with the heme cofactor shown as red sticks: cytochrome P450<sub>BM3</sub> (PDB 2IJ2), sperm whale myoglobin (PDB 1A6K), and *Rma* cytochrome *c* (PDB 3CP5). Reaction conditions were heme protein in *E. coli* whole cells (re-suspended to OD<sub>600</sub> = 30), 10 mM substrate **51a**, 10 mM ethyl diazoacetate, 5 vol% EtOH in M9-N buffer at room temperature under anaerobic conditions for 18 hours. TTN, total turnover number; N.D., not detected; WT, wild type; Mb, sperm whale myoglobin, HGG, Hell's Gate globin; cyt *c*, cytochrome *c*; *Hth*, *Hydrogenobacter thermophilus*.

With P411 P-4 A82L as the starting template, sequential rounds of site-saturation mutagenesis and screening in whole *E. coli* cells were performed to identify increasingly active and enantioselective biocatalysts for C–H alkylation (see Section C.4 in Appendix C for details). Amino acid residues chosen for mutagenesis included those which line the active site pocket, reside on loops and other flexible regions of the protein, or possess a nucleophilic side chain.<sup>26</sup> Improved variants were subsequently evaluated in reactions using clarified *E. coli* lysate with *p*-methoxybenzyl methyl ether (**51a**) and 4-ethylanisole.

Five rounds of mutagenesis and screening yielded variant P411-gen6, which furnished product **52a** with 60 TTN. Unlike the native monooxygenase activity, the C–H alkylation process does not require reducing equivalents from the FAD and FMN domains of the enzyme. Surmising that these domains may not be needed for the C–H alkylation reaction, we performed systematic truncations of P411-gen6 to determine the minimally sufficient domain(s) for retaining catalytic activity (Figure 4-3). Curiously, removal of the FAD domain, containing 37% of the amino acids in the full-length protein, created an enzyme with higher C–H alkylation activity: P411ΔFAD-gen6 delivers **52a** with 100 TTN, a 1.7-fold increase in TTN compared with P411-gen6. This indicates that the FAD domain may have (negative) allosteric effects on C–H alkylation activity. Further studies with these

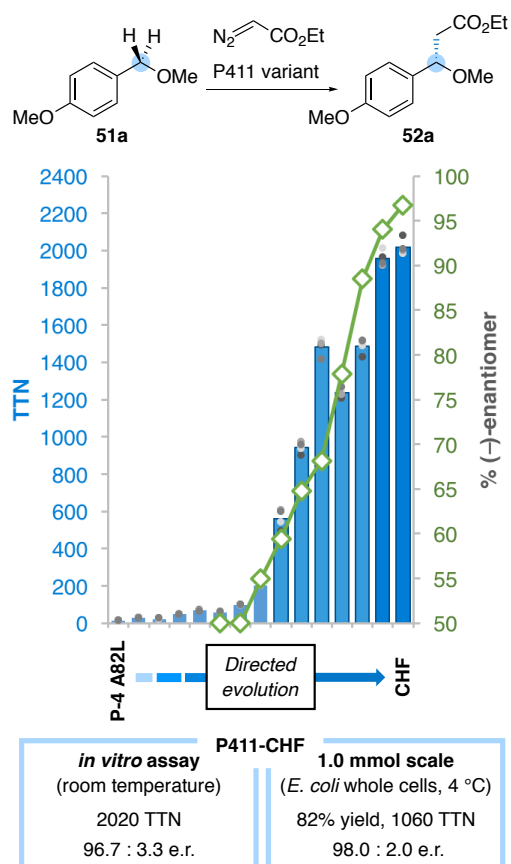
truncated enzymes revealed that they could be used in whole *E. coli* cells, in clarified *E. coli* cell lysate, and as purified proteins (Table C-3 in Appendix C).



**Figure 4-3.** Truncation of a full-length P411 protein delivers active C–H alkylation enzymes. **(a)** Domain architecture of cytochrome P450<sub>BM3</sub>. For its native monooxygenase activity, the FMN and FAD domains, collectively called the reductase domain, are responsible for delivering the necessary reducing equivalents from NADPH to the heme domain. The end of the FMN domain and the fragment of the polypeptide chain included in the  $\Delta$ FAD complex were chosen based on a report by S. Govindaraj and T. L. Poulos.<sup>27</sup> **(b)** Systematic truncation of the P411-gen6 full-length protein was performed to deliver P411-gen6b (P411 $\Delta$ FAD-gen6, amino acids 1–664) and P411-gen6 heme-domain only (amino acids 1–463). Standard reaction conditions: lysate of *E. coli* with 2.0  $\mu$ M heme protein, 10 mM **51a**, 10 mM ethyl diazoacetate, and 1 mM Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub> (unless otherwise indicated). TTN results are an average of at least duplicate reactions. RT, room temperature; TTN, total turnover number. <sup>†</sup>5 mM dithionite was used in these reactions.

Eight additional rounds of mutagenesis and screening yielded P411-CHF (P411 $\Delta$ FAD C–H Functionalization enzyme, full list of changes provided in the Supplementary Information). The activities of the enzymes developed over the course of this directed evolution evaluated on substrates **51a** and ethyl diazoacetate are summarized in Figure 4-4 (see Figure C-1 in Appendix C for results with 4-ethylanisole and ethyldiazoacetate). P411-CHF displays 140-fold improvement in activity over P-4 A82L and delivers **52a** with excellent stereoselectivity (2020 TTN, 96.7 : 3.3 e.r. using clarified *E. coli* lysate).

Subsequent studies showed that the stereoselectivity could be improved by conducting the reaction at lower temperature (e.g. 4 °C) with no significant change to TTN (Table C-4 in Appendix C). Enzymatic C–H alkylation can be performed on millimole scale: using 1.0 mmol substrate **51a**, *E. coli* harboring P411-CHF at 4 °C furnished **52a** in 82% isolated yield, 1060 TTN, and 98.0 : 2.0 e.r. (Figure 4-4).

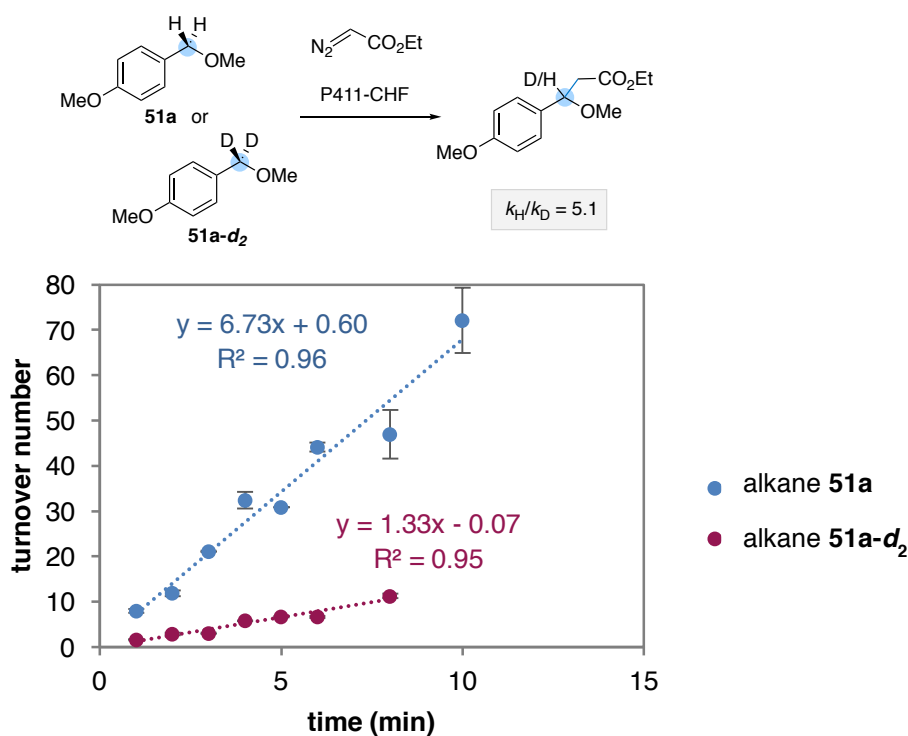


**Figure 4-4.** Directed evolution of a cytochrome P411 for enantioselective C–H alkylation. Bars represent mean TTN values averaged over four reactions (performed from two independent cell cultures; each used for duplicate reactions); each TTN data point is shown as a grey dot. Enantioselectivity results are represented by green diamonds. Reaction conditions were P411 variant in clarified *E. coli* lysate, 1 mM  $\text{Na}_2\text{S}_2\text{O}_4$ , 10 mM substrate **51a**, 10 mM ethyl diazoacetate, 5 vol% EtOH in M9-N buffer at room temperature under anaerobic conditions for 18 hours.

#### 4.3.2 Kinetic isotope effect study

Preliminary mechanistic investigations were pursued to interrogate the nature of the C–H insertion step. Independent initial rates measured for reactions with substrate **51a** or deuterated substrate **51a- $d_2$**  revealed a normal kinetic isotope effect (KIE,  $k_{\text{H}}/k_{\text{D}}$ ) of 5.1 for C–H alkylation catalyzed by P411-CHF (Figure 4-5). This suggests that C–H insertion is rate-determining and could possibly involve a linear transition state<sup>28</sup>. In contrast, kinetic

isotope effects for rhodium catalysts with carboxylate ligands are significantly less ( $\text{KIE} = 1.55\text{--}3.2$ )<sup>29,30</sup>; this has been invoked as evidence to support a widely accepted three-centered transition state for C–H insertion with these systems<sup>31</sup>. (*Note:* KIE values for rhodium-carboxylate catalysts are commonly measured using an intermolecular competition experiment.) The difference in KIE between P411-CHF and the rhodium-carboxylate catalysts suggests that these systems may have different transition states or different mechanisms for the C–H insertion step. Since the nature of the C–H insertion step could influence the substrate and product profiles of the catalyst, this is one strong motivation to develop diverse systems for this chemistry.



**Figure 4-5.** Kinetic isotope effect of C–H alkylation catalyzed by P411-CHF. Data points represent an average of duplicate measurements; error bars represent one standard deviation. Data collected at the 10-minute time point using substrate **51a-d<sub>2</sub>** were excluded due to non-linear behavior. Detailed experimental methods are described in Section 5.5 Experimental Methods.

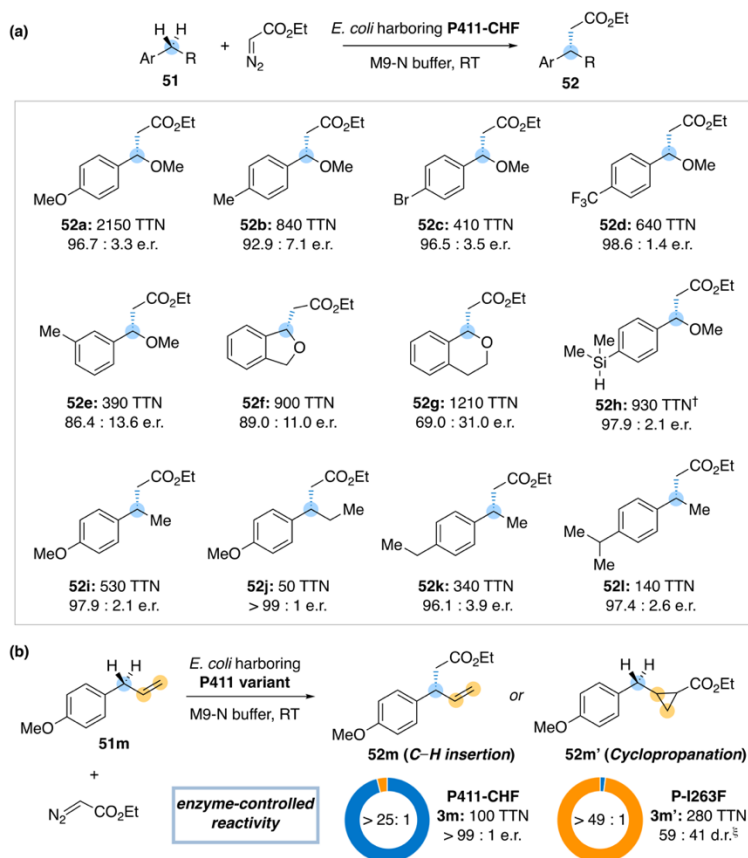


### 4.3.3 Benzylic C–H insertion substrate scope

Using *E. coli* harboring P411-CHF, we assayed a range of benzylic substrates for coupling with ethyl diazoacetate (Figure 4-6). Both electron-rich and electron-deficient functionalities on the aromatic ring are well-tolerated (**52a–52e**, **52h**); cyclic substrates are also suitable coupling partners (**52f**, **52g**). Functionalization of alkyl benzenes is successful at secondary benzylic  $sp^3$  C–H bonds (**52i–52l**). Notably, in the biotransformation of substrate **51l** containing both tertiary and secondary benzylic C–H bonds, P411-CHF preferentially functionalizes the secondary position despite its higher C–H bond dissociation energy (BDE). The carbene intermediate derived from ethyl diazoacetate belongs to the acceptor-only class. In contrast to the more widely-used donor/acceptor carbenes, acceptor-only intermediates are more electrophilic, and as a result selective reactions with this carbene class are still a major challenge for small-molecule catalysts.<sup>15,19</sup> Our results show that P411-CHF can control this highly reactive intermediate to furnish the desired  $sp^3$  C–H alkylation products and do so with high enantioselectivity.

Enzymes can exhibit excellent reaction selectivity arising from their ability to form multiple interactions with substrates and intermediates throughout a reaction cycle. We hypothesized that the protein scaffold could be tuned to create complementary enzymes which access different reaction outcomes available to a substrate. When P411-CHF was challenged with 4-allylanisole (**51m**), a substrate which can undergo both C–H alkylation and cyclopropanation, we observed that C–H alkylation product **52m** dominates, with selectivity > 25:1 (Figure 4-6, Figure C-4 in Appendix C). In contrast, a related full-length P411 variant P-I263F, containing thirteen mutations in the heme domain relative to P411-CHF, catalyzed only the formation of cyclopropane product **52m'**. Additionally, despite the established reactivity of silanes with iron-carbene,<sup>10</sup> P411-CHF delivered C–H alkylation product **52h** when substrate **51h** was employed in the reaction (Si–H insertion product **52h'** was also observed but its formation may not be catalyzed by P411-CHF, Figure C-5 in Appendix C). Reaction with P-I263F, in contrast, provided only the Si–H insertion product. These examples demonstrate an exceptional feature of macromolecular

enzymes: different products can be obtained simply by changing the amino acid sequence of the protein catalyst.



**Figure 4-6.** Substrate scope for benzylic C–H alkylation with P411-CHF. **(a)** Experiments were performed using *E. coli* expressing cytochrome P411-CHF (OD<sub>600</sub> = 30) with 10 mM substrate **51a–52l** and 10 mM ethyl diazoacetate at room temperature (RT) under anaerobic conditions for 18 hours; reported TTNs are the average of four reactions (performed from two independent cell cultures, each used for duplicate reactions). <sup>†</sup>Si–H insertion product **52h'** is also observed (see Figure C-5 in Appendix C). **(b)** Reaction selectivity for carbene C–H insertion or cyclopropanation can be controlled by the protein scaffold. Experiments were performed as in **(a)** using the indicated P411 variant. <sup>§</sup>d.r. is given as *cis* : *trans*; e.r. was not determined.

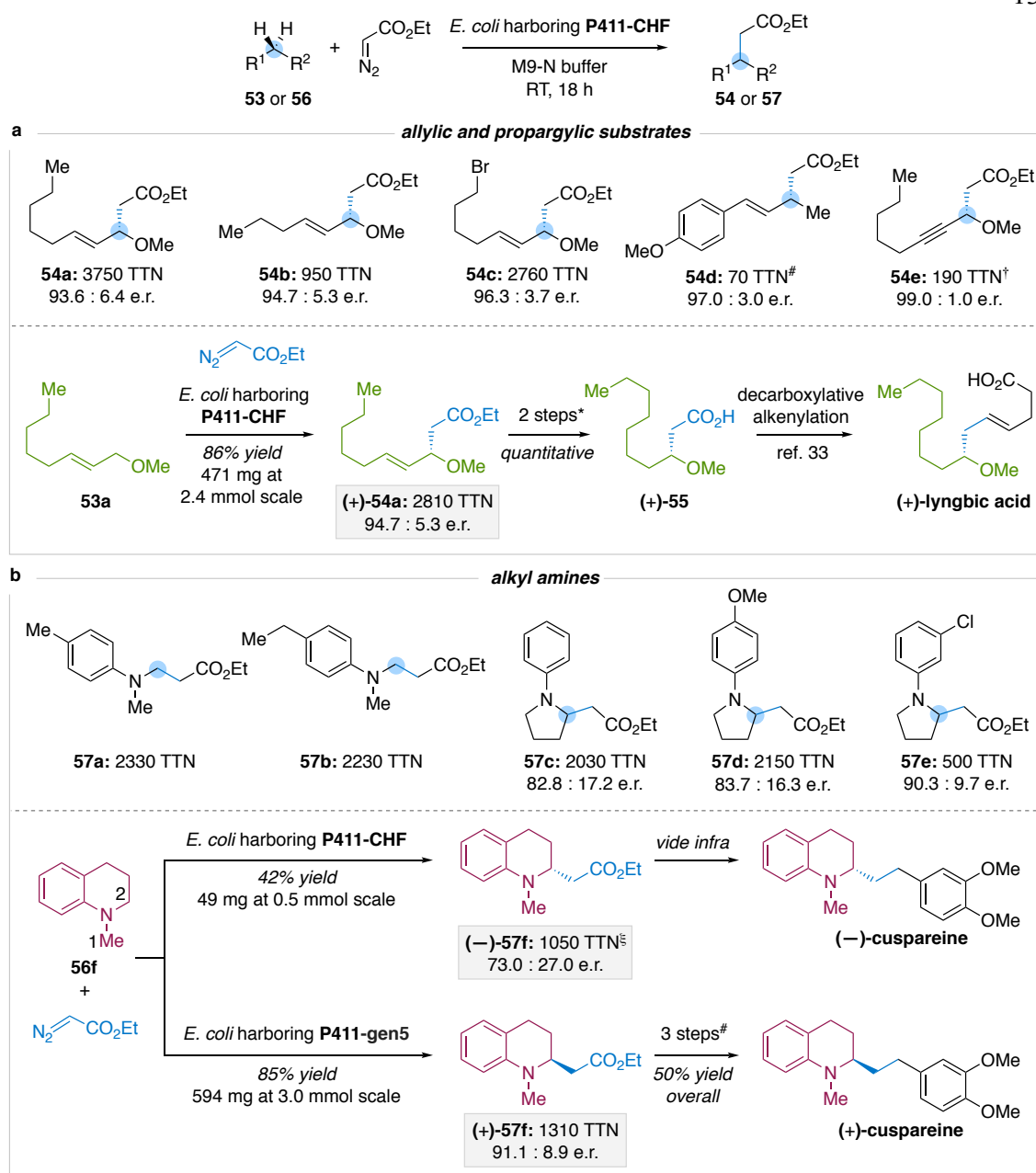
#### 4.3.4 Enzymatic functionalization of allylic, propargylic, and $\alpha$ -amino C–H bonds

Enzymatic C–H alkylation is not limited to functionalization of benzylic C–H bonds. Structurally dissimilar molecules containing allylic or propargylic C–H bonds are excellent substrates for this chemistry (Figure 4-7a). In contrast to **51a–51m**, which contain a rigid

benzene ring, compounds **53a–53c** and **53e** feature flexible linear alkyl chains. Their successful enantioselective alkylation suggests that the enzyme active site can accommodate substrate conformational flexibility while enforcing a favoured substrate orientation relative to the carbene intermediate. To demonstrate the utility of this biotransformation, we applied the methodology to the formal synthesis of lyngbic acid (Figure 4-7a). Marine cyanobacteria incorporate this versatile biomolecule into members of the malyngamide family of natural products; likewise, total synthesis approaches to these natural products typically access lyngbic acid as a strategic intermediate *en route* to the target molecules.<sup>32</sup> Using *E. coli* harboring P411-CHF, intermediate **54a** was produced on 2.4 mmol scale in 86% isolated yield, 2810 TTN, and 94.7 : 5.3 e.r.. Subsequent hydrogenation and hydrolysis provided (*R*)-(+)-**55** in quantitative yield, which can be elaborated to (*R*)-(+)-lyngbic acid by decarboxylative alkenylation.<sup>33</sup>

As part of our substrate scope studies, we challenged P411-CHF with alkyl amine compounds. Compounds of this type are typically challenging for C–H functionalization methods because the amine functionality may coordinate to and inhibit the catalyst or create the opportunity for undesirable side reactions (e.g. ylide formation and its associated rearrangements).<sup>34,35</sup> Using **56a** or **56b**, substrates which have both benzylic C–H bonds and  $\alpha$ -amino C–H bonds, P411-CHF delivered the corresponding  $\beta$ -amino ester product with high efficiency (**57a** and **57b**, Figure 4-7b). Notably, benzylic C–H insertion was not observed (with **56a**, Figure C-7 in Appendix C) or significantly suppressed (with **56b**, Figure C-8 in Appendix C), despite the typically lower BDEs of benzylic C–H bonds compared to  $\alpha$ -amino C–H bonds.<sup>36</sup> Additionally, *N*-aryl pyrrolidines (**56c–56e**) served as excellent substrates and were selectively alkylated at the  $\alpha$ -amino  $sp^3$  position. Using P411-CHF, the  $sp^3$  C–H alkylation of **56c** outcompetes a Friedel-Crafts type reaction on the aryl ring, which is a favourable process with other carbene-transfer systems.<sup>37,38</sup> Furthermore, alkylation product **57d** offers a conceivable strategy for the synthesis of  $\beta$ -homoproline, a motif which has been investigated for medicinal chemistry applications.<sup>39</sup>

Given that P411-CHF alkylates both primary and secondary  $\alpha$ -amino C–H bonds, we interrogated whether the enzyme could be selective for one of these positions. Employing *N*-methyl tetrahydroquinoline **56f** as the alkane substrate, P411-CHF afforded  $\beta$ -amino ester products with 1050 TTN and a 9 : 1 ratio of regioisomers (C2 : C1, and 73.0 : 27.0 e.r. for (–)-**57f**) (Figure 4-7b). As the tetrahydroquinoline ring is a privileged structural motif in natural products and bioactive molecules,<sup>40</sup> its selective functionalization could provide a concise strategy for the synthesis of alkaloids. To improve the selectivity for alkylation of **56f**, we tested variants along the evolutionary lineage from P-4 A82L to P411-CHF. We found that P411-gen5 had even better regioselectivity and delivered product with the opposite stereo-preference. In a 3.0 mmol scale reaction, *E. coli* harboring P411-gen5 delivered (+)-**57f** in 85% yield with excellent selectivity (1310 TTN, > 50 : 1 r.r., 91.1 : 8.9 e.r.). In only a few steps, the enzymatic product was successfully transformed to alkaloid (*R*)-(+)-cuspareine (Figure 4-7b).<sup>40</sup>

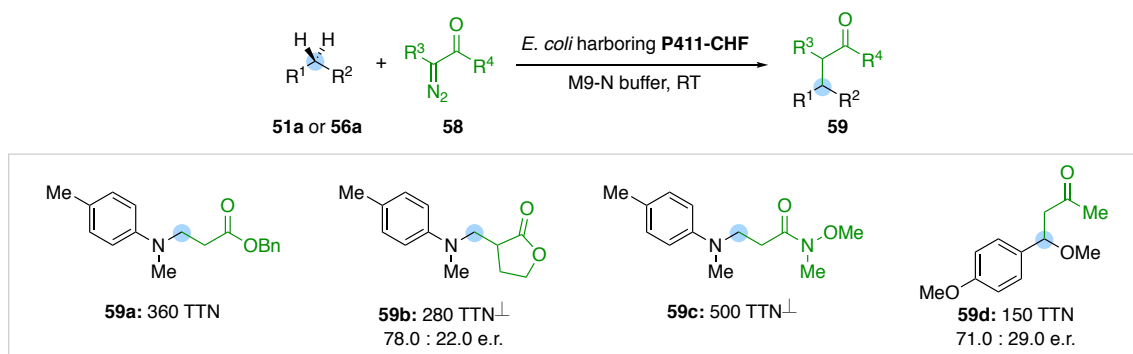


**Figure 4-7.** Application of P411 enzymes for  $sp^3$  C–H alkylation. **(a)** Allylic and propargylic C–H alkylation. Unless otherwise indicated, experiments were performed using *E. coli* expressing cytochrome P411-CHF with 10 mM substrate **53a–53e** and 10 mM ethyl diazoacetate; reported TTNs are the average of four reactions (performed from two independent cell cultures, each used for duplicate reactions). <sup>#</sup>TTN was calculated based on isolated yield from a reaction performed at 0.25 mmol scale. <sup>†</sup>Cyclopropene product was also observed (Figure C-6 in Appendix C). <sup>\*</sup>Hydrogenation, followed by hydrolysis. **(b)** Enzymatic alkylation of substrates containing  $\alpha$ -amino C–H bonds. Unless otherwise indicated, experiments were performed at 0.5 mmol scale using *E. coli* expressing

cytochrome P411-CHF with substrates **56a–56f** and ethyl diazoacetate; TTNs were calculated based on isolated yields of products shown. <sup>Ξ</sup>Isolated in 9 : 1 r.r. for **57f** : **57f'**.  
<sup>Ψ</sup>Reduction, halogen exchange, and Suzuki-Miyaura cross-coupling.

#### 4.3.5 Diazo reagent scope

We also probed the ability of select P411 enzymes to use other diazo reagents for C–H functionalization. Using different diazo reagents, enzymatic C–H alkylation can diversify one alkane substrate, such as **56a**, to several products (**59a–59c** in Figure 4-8 and Figure C-9 in Appendix C). The diazo substrate scope extends beyond ester-based reagents: Weinreb amide diazo compound **58c** and diazoketone **58d** were found to participate in enzymatic C–H alkylation to furnish products **59c** and **59d**, respectively. Additional substitution at the  $\alpha$ -position of the carbene, however, is generally not well-tolerated by P411-CHF and current related enzymes. With the exception of **58b**, reactions using disubstituted carbene reagents failed to yield appreciable amounts of desired products (Figure C-9 in Appendix C).



**Figure 4-8.** Enzymatic C–H alkylation with alternative diazo reagents. Unless otherwise indicated, reactions were performed at 0.5 mmol scale using *E. coli* expressing cytochrome P411-CHF with coupling partner **51a** or **56a** and diazo compounds **58a–58d**; TTNs were calculated based on isolated yields of products shown. <sup>⊥</sup>Variant P411-IY T327I was used. See Appendix C for the complete list substrates (Figure C-10 and Figure C-11) and information about enzyme variants.

## 4.4 Conclusion

This study demonstrates that a cytochrome P450 can acquire the ability to construct C–C bonds from  $sp^3$  C–H bonds and that activity and selectivity can be greatly enhanced using directed evolution. Nature provides a huge collection of possible alternative starting points for expanding the scope of this reaction even further and for achieving other selectivities. The cytochrome P450 superfamily can access an immense set of organic molecules for its native oxygenation chemistry; we envision that P411-derived enzymes and other natural heme protein diversity can be leveraged to generate families of C–H alkylation enzymes that emulate the scope and selectivity of nature's C–H oxygenation catalysts.

## 4.5 Experimental Methods

See Appendix C for supporting tables and figures, details about protein variants, synthesis and characterization of compounds, details regarding calibration curves, preparative scale enzymatic reactions, syntheses of (+)-lyngbic acid and (+)-cuspareine, and methods for determining enantioselectivity. Additional information, including nucleotide and amino acid sequences of variants, calibration curves and data analysis, chromatography traces, and  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra, can be found in the Supporting Information of the published paper.

### 4.5.1 General information

Unless otherwise noted, all chemicals and reagents were obtained from commercial suppliers (Sigma-Aldrich, VWR, Alfa Aesar, Combi-Blocks) and used without further purification. Silica gel chromatography was carried out using AMD Silica Gel 60, 230-400 mesh.  $^1\text{H}$  and  $^{13}\text{C}$  NMR spectra were recorded on a Varian Inova 300 MHz, Varian Inova 500 MHz, or Bruker Prodigy 400 MHz instrument in  $\text{CDCl}_3$  and are referenced to residual protio solvent signals. Data for  $^1\text{H}$  NMR are reported as follows: chemical shift ( $\delta$  ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, p = pentet, sext = sextet, m =

multiplet, dd = doublet of doublets, dt = doublet of triplets, ddd = doublet of doublet of doublets), coupling constant (Hz), integration. Sonication was performed using a Qsonica Q500 sonicator. High-resolution mass spectra were obtained at the California Institute of Technology Mass Spectral Facility. Synthetic reactions were monitored by thin layer chromatography (TLC, Merck 60 gel plates) using a UV-lamp or an appropriate TLC stain for visualization.

*E. coli* cells were grown using Luria-Bertani medium (LB) or Hyperbroth (AthenaES) (HB) with 0.1 mg/mL ampicillin (LB<sub>amp</sub> or HB<sub>amp</sub>). Primer sequences are available upon request. T5 exonuclease, Phusion polymerase, and Taq ligase were purchased from New England Biolabs (NEB, Ipswich, MA). M9-N minimal medium (abbreviated as M9-N buffer; pH 7.4) was used as a buffering system for whole cells and lysates, unless otherwise specified. M9-N buffer was used without a nitrogen source; it contains 47.7 mM Na<sub>2</sub>HPO<sub>4</sub>, 22.0 mM KH<sub>2</sub>PO<sub>4</sub>, 8.6 mM NaCl, 2.0 mM MgSO<sub>4</sub>, and 0.1 mM CaCl<sub>2</sub>.

#### 4.5.2 Chromatography

Analytical reversed-phase high-performance liquid chromatography (HPLC) was carried out using an Agilent 1200 series instrument and a Kromasil 100 C18 column (4.6 × 50 mm, 5 μm) or an Eclipse XDB C18 column (4.6 × 150 mm, 3 μm) with water and acetonitrile as the mobile phase. Analytical chiral HPLC was conducted using either an Agilent 1200 series instrument with *n*-hexane and isopropanol as the mobile phase or JACSO 2000 series supercritical fluid chromatography (SFC) system with supercritical CO<sub>2</sub> and isopropanol as the mobile phase. Enantiomers were separated using one of the following chiral columns: Chiralpak AD-H, Chiralpak IC (4.6 mm × 25 cm), Chiralcel OB-H (4.6 mm × 25 cm), Chiralcel OD-H (4.6 mm × 25 cm), Chiralcel OJ-H (4.6 mm × 25 cm). Gas chromatography (GC) analysis was carried out using an Agilent 7820A or Shimadzu GC-17A GC system, both equipped with an FID detector and with a J&W HP-5 column (30 m × 0.32 mm, 0.25 μm film). Chiral GC was conducted using either an Agilent 7820A instrument (FID) and an Agilent CycloSil-B column (30 m × 0.32 mm, 0.25 μm film) or an Agilent 6850 GC (FID) with a Chiraldex G-TA column (30 m × 0.25 mm,



0.12  $\mu\text{m}$  film). Gas chromatography-mass spectrometry (GC-MS) analyses were carried out using a Shimadzu GCMS-QP2010SE system and J&W HP-5ms column (30 m  $\times$  0.25 mm, 0.25  $\mu\text{m}$  film).

#### 4.5.3 Cloning and site-saturation mutagenesis

pET22b(+) was used as a cloning and expression vector for all enzymes described in this study. All enzymes described in this study were expressed with a C-terminal 6xHis-tag. Site-saturation mutagenesis was performed using the “22c-trick” method.<sup>41</sup> The PCR products were digested with *DpnI*, gel purified, and ligated using Gibson Mix<sup>TM</sup> (ref. 42). The ligation mixture was used to directly transform electrocompetent *E. coli* strain *E. cloni* BL21(DE3) cells (Lucigen).

#### 4.5.4 Expression of P450 and P411 variants in 96-well plates

Single colonies from LB<sub>amp</sub> agar plates were picked using sterile toothpicks and cultured in deep-well 96-well plates containing LB<sub>amp</sub> (300  $\mu\text{L}$ /well) at 37 °C, 220 rpm shaking, and 80% relative humidity overnight. After, HB<sub>amp</sub> (1000  $\mu\text{L}$ /well) in a deep-well 96-well plate was inoculated with an aliquot (50  $\mu\text{L}$ / well) of these overnight cultures and allowed to shake for 3 hours at 37 °C, 220 rpm, and 80% relative humidity. The plates were cooled on ice for 30 minutes and the cultures were induced with 0.5 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) and 1.0 mM 5-aminolevulinic acid (final concentrations). Expression was conducted at 20 °C, 150 rpm for 16–20 hours.

#### 4.5.5 Reaction screening in 96-well plate format

*E. coli* (*E. cloni* BL21(DE3)) cells in deep-well 96-well plates were pelleted (3,000  $\times$  g, 5 min, RT) and resuspended in M9-N buffer (20  $\mu\text{L}$ /well) by gentle vortexing. A GOX oxygen depletion system was added (20  $\mu\text{L}$ /well of a stock solution containing 14,000 U/mL catalase and 1,000 U/mL glucose oxidase in 0.1 M potassium phosphate buffer, pH 8.0) and the 96-well plate was transferred into an anaerobic chamber. In the anaerobic chamber, argon-sparged reaction buffer (50 mM glucose in M9-N or 33 mM glucose in

M9-N, 300  $\mu\text{L}$ /well) was added, followed by alkane substrate (10  $\mu\text{L}$ /well, 400 mM in EtOH) and ethyl diazoacetate (10  $\mu\text{L}$ /well, 400 mM in EtOH). In some cases, the substrates and reaction buffer were mixed together prior to addition to the plate. The plate was sealed with an aluminum foil and shaken at room temperature and 500 rpm in the anaerobic chamber. After 5–20 hours, the seal was removed and the reactions were worked up following the appropriate method below.

*Product formation screening using GC and GC-MS.* After 5–20 hours, a solution of 0.4 mM 1,3,5-trimethoxybenzene (internal standard) in a mixed solvent system (cyclohexane/ ethyl acetate = 1:1, 510  $\mu\text{L}$ ) was added. The plate was tightly sealed with a reusable silicone mat, vortexed (15 s  $\times$  3), and centrifuged (3,000  $\times$  g, 5 min) to completely separate the organic and aqueous layers. The organic layers (180  $\mu\text{L}$ /well) were transferred to 300  $\mu\text{L}$  vial inserts, which were then placed in 2 mL vials and analyzed by GC.

*Product formation screening using HPLC.* After 5–20 hours, the reaction mixtures, or an aliquot thereof (150  $\mu\text{L}$ /well), were quenched by the addition of an equal or greater volume of acetonitrile (400  $\mu\text{L}$ /well or 150–200  $\mu\text{L}$ /well). This step is kept consistent within each round of directed evolution. The plate containing the resulting mixture was tightly sealed with a reusable silicone mat, vortexed (15 s  $\times$  3), and centrifuged (3,000  $\times$  g, 5 min) to pellet the cells. The supernatant was filtered through an AcroPrep 96-well filter plate (0.2  $\mu\text{m}$ ) into a shallow-well plate and analyzed by reverse-phase HPLC.

*Enantioselectivity screening.* After 5–24 hours, mixed solvent (cyclohexane/ ethyl acetate = 1:1, 250–500  $\mu\text{L}$ / well) was added to the reaction mixtures or aliquots thereof (250  $\mu\text{L}$ ). The plate containing the resulting mixture was tightly sealed with a reusable silicone mat, vortexed (15 s  $\times$  3) and centrifuged (3,000  $\times$  g, 5 min) to completely separate the organic and aqueous layers. When smaller volumes of mixed solvent were used for the extraction (< 400  $\mu\text{L}$ ), the extraction mixture was transferred to a 1.6 mL Eppendorf tube, vortexed (15 s  $\times$  3), and centrifuged (20,000  $\times$  g, 1 min). The organic layers (180  $\mu\text{L}$ /well) were transferred to 300  $\mu\text{L}$  vial inserts, which were then placed in 2 mL vials and analyzed by chiral HPLC (IC column, 2% *i*-PrOH in *n*-hexane).

#### 4.5.6 Expression of P411 variants.

*E. coli* (*E. coli* BL21(DE3)) cells carrying plasmid encoding the appropriate P411 variant were grown overnight in 5 mL LB<sub>amp</sub>. Preculture (2 mL) was used to inoculate 48 mL of HB<sub>amp</sub> in a 125 mL Erlenmeyer flask; this culture was incubated at 37 °C, 230 rpm for 2.5 hours. The culture was then cooled on ice (20–30 min) and induced with 0.5 mM IPTG and 1.0 mM 5-aminolevulinic acid (final concentrations). Expression was conducted at 20 °C, 130 rpm, for 16–18 hours. Following, *E. coli* cells were pelleted by centrifugation ( $2,600 \times g$ , 10 min, 4 °C or  $3,000 \times g$ , 5 min, 4 °C). Media was removed and the resulting cell pellet was resuspended in M9-N buffer to OD<sub>600</sub> = 60. An aliquot of this cell suspension (3 mL) was taken to determine P411 concentration using the hemochrome assay after lysis by sonication. When applicable, remaining cell suspension was further diluted with M9-N buffer to the OD<sub>600</sub> used for the biotransformation (described in (4.5.8)) and the concentration of P411 protein in the biotransformation was calculated accordingly.

#### 4.5.7 Hemochrome assay for the determination of heme protein concentration.

*E. coli* cells expressing heme protein and resuspended in M9-N buffer were lysed by sonication using a Qsonica Q500 sonicator equipped with a microtip (2 mins, 1 second on, 1 second off, 25% amplitude); samples were kept on wet ice for this process. The resulting lysed solution was centrifuged ( $20,000 \times g$ , 10 min, 4 °C) to remove cell debris. The supernatant (clarified lysate) was separated from the pellet and kept on ice until use. In a falcon tube, a solution of 0.2 M NaOH, 40% (v/v) pyridine, 0.5 mM K<sub>3</sub>Fe(CN)<sub>6</sub> was prepared (pyridine-NaOH-K<sub>3</sub>Fe(CN)<sub>6</sub> solution). Separately, a solution of 0.5 M Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub> (sodium dithionite) was prepared in 0.1 M NaOH. To an Eppendorf tube containing 500 µL of clarified lysate in M9-N buffer was added 500 µL of the pyridine-NaOH-K<sub>3</sub>Fe(CN)<sub>6</sub> solution, mixed, and transferred to a cuvette; the UV-Vis spectrum of the oxidized Fe<sup>III</sup> state was recorded immediately. To the cuvette was then added 10 µL of the sodium dithionite solution. The cuvette was sealed with parafilm and the UV-Vis spectrum of the reduced Fe<sup>II</sup> state was recorded immediately. A cuvette containing 500 µL of M9-N, 100 µL 1 M NaOH, 200 µL pyridine, and 200 µL water (complete mixture without protein and

$\text{K}_3\text{Fe}(\text{CN})_6$ ) was used as a reference for all absorbance measurements. Concentrations of cytochromes P450, cytochromes P411, and globins were determined using a published extinction coefficient for heme *b*,  $\epsilon_{556(\text{reduced})-540(\text{oxidized})} = 23.98 \text{ mM}^{-1}\text{cm}^{-1}$  (ref. 43). Cytochrome *c* concentration was measured using a modified procedure, reported previously.<sup>10</sup>

#### 4.5.8 Biotransformations using whole *E. coli* cells

Suspensions of *E. coli* (*E. coli* BL21(DE3)) cells expressing the appropriate heme protein variant in M9-N buffer (typically  $\text{OD}_{600} = 30$ ) were degassed by bubbling with argon in sealed vials for at least 40 minutes; the cells were kept on ice during this time. Separately, a solution of *D*-glucose (250 mM in M9-N) was degassed by sparging with argon for at least 30 minutes. All solutions were then transferred into an anaerobic chamber for reaction set up. To a 2 mL vial were added a GOX oxygen depletion solution (20  $\mu\text{L}$  of stock solution containing 14,000 U/mL catalase and 1,000 U/mL glucose oxidase in 0.1 M potassium phosphate buffer, pH 8.0), *D*-glucose (40  $\mu\text{L}$  of 250 mM stock solution in M9-N buffer), degassed suspension of *E. coli* expressing P411 (typically  $\text{OD}_{600} = 30$ , 320  $\mu\text{L}$ ), alkane substrate (10  $\mu\text{L}$  of 400 mM stock solution in EtOH), and diazo compound (10  $\mu\text{L}$  of 400 mM stock solution in EtOH) in the listed order. Final reaction volume was 400  $\mu\text{L}$ ; final concentrations were 10 mM alkane substrate, 10 mM diazo compound, and 25 mM *D*-glucose. **Note:** reaction performed with *E. coli* cells resuspended to  $\text{OD}_{600} = 30$  indicates that 320  $\mu\text{L}$  of  $\text{OD}_{600} = 30$  cells were added, and likewise for other reaction  $\text{OD}_{600}$  descriptions. The vials were sealed, removed from the anaerobic chamber, and shaken at room temperature and 500 rpm for 18 hours. A modified procedure was used for reactions conducted at 4 °C. Reactions were set up in the same manner, except kept on ice. Reactions were shaken in a cold room (4 °C) and 500 rpm for 18 hours. The reactions were worked up and analyzed by HPLC or GC as appropriate. Preparative scale enzymatic reactions were performed using a different procedure which is described in detail in **Section C.8**. The expression of heme protein was measured using the hemochrome assay (**Section 4.5.7**), and the concentration of heme protein in the biotransformation was calculated

accordingly.

#### 4.5.9 Enzymatic reactions using clarified *E. coli* lysate.

Lysates for biocatalytic reactions were prepared as follows: *E. coli* (*E. coli* BL21(DE3)) cells expressing the appropriate heme protein variant were resuspended in M9-N buffer and adjusted to  $OD_{600} = 60$ . The cell suspension, in 3 mL portions, was lysed by sonication using a Qsonica Q500 sonicator equipped with a microtip (2 mins, 1 second on, 1 second off, 25% amplitude); samples were kept on wet ice for this process. The resulting lysed solution was centrifuged ( $20,000 \times g$ , 10 min, 4 °C) to remove cell debris. Protein concentration of the supernatant (clarified lysate) was determined using the hemochrome assay (see **Section 4.5.7**); the protein concentration in lysate was adjusted to the desired amount by addition of M9-N buffer. Lysate was placed in a sealed vial and the headspace of the vial was purged with a stream of argon for at least 40 minutes. The lysate was kept on ice during all parts of this procedure. Separately, *D*-glucose solution (500 mM in M9-N buffer) and  $Na_2S_2O_4$  (20 mM in M9-N) were degassed by bubbling the solutions with argon for at least 40 minutes. All solutions were then transferred into an anaerobic chamber for reaction set up. To a 2 mL vial were added a GOX oxygen depletion solution (20  $\mu$ L of stock solution containing 14,000 U/mL catalase and 1,000 U/mL glucose oxidase in 0.1 M potassium phosphate buffer, pH 8.0), *D*-glucose (20  $\mu$ L of 500 mM stock solution in M9-N buffer), lysate (320  $\mu$ L),  $Na_2S_2O_4$  (20  $\mu$ L of 20 mM solution in M9-N), alkane substrate (10  $\mu$ L of 400 mM stock solution in EtOH), and ethyl diazoacetate (10  $\mu$ L of 400 mM stock solution in EtOH) in the listed order. Final reaction volume was 400  $\mu$ L; final concentrations were typically 2.0  $\mu$ M heme protein, 1 mM  $Na_2S_2O_4$ , 10 mM alkane substrate, 10 mM ethyl diazoacetate, and 25 mM *D*-glucose. The vials were sealed, removed from the anaerobic chamber, and shaken at room temperature and 500 rpm for 18 hours. Reactions were analyzed following the same methods as described for biotransformations using whole *E. coli* cells.

#### 4.5.10 Protein purification

*E. coli* (*E. coli* BL21(DE3)) cells carrying plasmid encoding a P411 variant were grown overnight in 19 mL LB<sub>amp</sub> (37 °C, 250 rpm). HB<sub>amp</sub> (450 mL) in a 1 liter flask was inoculated with 19 mL of the preculture and shaken for 2.5 hours at 37 °C and 220 rpm. Cultures were cooled on ice (30 min) and induced with 0.5 mM IPTG and 1.0 mM 5-aminolevulinic acid (final concentrations). Expression was conducted at 22 °C, 130 rpm, for 16–20 hours. Cultures were then centrifuged ( $5,000 \times g$ , 10 min, 4 °C) and the cell pellets frozen at -20 °C. For purification, frozen cells from two such cultures were combined; purification was conducted according to the previously described procedure of **Section 3.5.7**. Concentrated proteins were aliquoted, flash-frozen on powdered dry ice, and stored at -80 °C until use.

#### **4.5.11 Enzymatic reactions using purified heme proteins**

A solution of NADPH or Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub> in phosphate buffer (0.1 M potassium phosphate, pH 8.0) was degassed by sparging with argon for at least 30 minutes (reaction solution). Separately, a solution of *D*-glucose (250 mM in 0.1 M potassium phosphate, pH 8.0) was also degassed in the same manner. Crimp vials (2 mL) were each charged with the GOX oxygen depletion system (20 µL of a stock solution containing 14,000 U/mL catalase and 1,000 U/mL glucose oxidase in 0.1 M potassium phosphate, pH 8.0). After degassing was complete, all solutions, 2 mL crimp vials, and purified protein (in 0.1 M potassium phosphate, pH 8.0), kept on ice, were brought into the anaerobic chamber. *D*-Glucose (40 µL of 250 mM solution), reaction buffer containing reductant (300 µL, reductant is either NADPH or Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub>), and purified protein (20 µL) were added to the 2 mL vial. The mixture was shaken for 5 min to ensure even distribution of the proteins; then, reaction vials were charged with alkane substrate (10 µL of 400 mM stock solution in EtOH) and ethyl diazoacetate (10 µL of 400 mM stock solution in EtOH). Final reaction volume was 400 µL. Final concentrations were 10 mM alkane substrate, 10 mM ethyl diazoacetate, and 25 mM *D*-glucose; concentrations of protein and reductant are variable and described with the individual experiment. The vials were sealed, removed from the anaerobic chamber, and shaken at room temperature and 500 rpm for 18 hours. The reactions were quenched

by the addition of acetonitrile (400  $\mu$ L) and internal standard (10  $\mu$ L). This mixture was then transferred to a microcentrifuge tube and centrifuged (20,000  $\times$  g, 10 min). The supernatant was transferred to a vial and analyzed by HPLC.

#### 4.6 References and Notes

---

1. Hartwig, J. F.; Larsen, M. A. Undirected, homogeneous C–H bond functionalization: Challenges and opportunities. *ACS Cent. Sci.* **2016**, *2*, 281–292.
2. Saint-Denis, T. G.; Zhu, R.-Y.; Chen, G.; Wu, Q.-F.; Yu, J.-Q. Enantioselective C(sp<sup>3</sup>)–H bond activation by chiral transition metal catalysts. *Science* **2018**, *359*, doi: 10.1126/science.aao4798.
3. Frey, P. A.; Hegeman, A. D. *Enzymatic Reaction Mechanisms* (Oxford University Press, New York, **2007**), chap. 14.
4. Yokoyama, K.; Lilla, E. A. C–C bond forming radical SAM enzymes involved in the construction of carbon skeletons of cofactors and natural products. *Nat. Prod. Rep.* **2018**, *35*, 660–694.
5. Bauerle, M. R.; Schwalm, E. L.; Booker, S. J. Mechanistic diversity of radical S-adenosylmethionine (SAM)-dependent methylation. *J. Biol. Chem.* **2015**, *290*, 3995–4002.
6. McLaughlin, M. I.; van der Donk, W. A. Stereospecific radical-mediated B12-dependent methyl transfer by the fosfomycin biosynthesis enzyme Fom3. *Biochemistry* **2018**, *57*, 4967–4971.
7. Shisler, K. A.; Broderick, J. B. Glycyl radical activating enzymes: Structure, mechanism, and substrate interactions. *Arch. Biochem. Biophys.* **2014**, *546*, 64–71.
8. Poulos, T. L. Heme enzyme structure and function. *Chem. Rev.* **2014**, *114*, 3919–3962.
9. Coelho, P. S.; Brustad, E. M.; Kannan, A.; Arnold, F. H. Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes. *Science* **2013**, *339*, 307–310.
10. Kan, S. B. J.; Lewis, R. D.; Chen, K.; Arnold, F. H. Directed evolution of cytochrome *c* for carbon–silicon bond formation: Bringing silicon to life. *Science* **2016**, *354*, 1048–1051.

11. Brandenberg, O. F.; Fasan, R.; Arnold, F. H. Exploiting and engineering hemoproteins for abiological carbene and nitrene transfer reactions. *Curr. Opin. Biotechnol.* **2017**, *47*, 102–111.
13. Vargas, D.; Tinoco, A.; Tyagi, V.; Fasan, R. Myoglobin-catalyzed C–H functionalization of unprotected indoles. *Angew. Chem. Int. Ed.* **2018**, *57*, 9911–9915.
14. Liao, K.; Pickel, T. C.; Boyarskikh, V.; Bacsa, J.; Musaev, D. G.; Davies, H. M. L. Site-selective and stereoselective functionalization of non-activated tertiary C–H bonds. *Nature* **2017**, *551*, 609–613.
15. Weldy, N. M.; Schafer, A. G.; Owens, C. P.; Herting, C. J.; Varela-Alvarez, A.; Chen, S.; Niemeyer, Z.; Musaev, D. G.; Sigman, M. S.; Davies, H. M. L.; Blakey, S. B. Iridium(III)-bis(imidazolynyl)phenyl catalysts for enantioselective C–H functionalization with ethyl diazoacetate. *Chem. Sci.* **2016**, *7*, 3142–3146.
17. Wang, Y.; Wen, X.; Cui, X.; Zhang, X. P. Enantioselective radical cyclization for construction of 5-membered ring structures by metalloradical C–H alkylation. *J. Am. Chem. Soc.* **2018**, *140*, 4792–4796.
18. Caballero, A.; Despagnet-Ayoub, E.; Díaz-Requejo, M. M.; Díaz-Rodríguez, A.; González-Núñez, M. E.; Mello, R.; Muñoz, B. K.; Ojo, W.-S.; Asensio, G.; Etienne, M.; Pérez, P. J. Silver-catalyzed C–C bond formation between methane and ethyl diazoacetate in supercritical CO<sub>2</sub>. *Science* **2011**, *332*, 835–838.
19. Wu, W.-T.; Yang, Z.-P.; You, S.-L. in *Asymmetric Functionalization of C–H Bonds*, ed. You, S.-L. (Royal Society of Chemistry, Cambridge, UK, **2015**), chap. 1.
20. Key, H. M.; Dydio, P.; Clark, D. S.; Hartwig, J. F. Abiological catalysis by artificial haem proteins containing noble metals in place of iron. *Nature* **2016**, *534*, 534–537.
21. Mbuvi, H. M.; Woo, L. K. Catalytic C–H insertions using iron(III) porphyrin complexes. *Organometallics* **2008**, *27*, 637–645.
22. Li, Y.; Huang, J.-S.; Zhou, Z.-Y.; Che, C.-M.; You, X.-Z. Remarkably stable iron porphyrins bearing nonheteroatom-stabilized carbene or (alkoxycarbonyl)carbenes: Isolation, X-ray crystal structures, and carbon atom transfer reactions with hydrocarbons. *J. Am. Chem. Soc.* **2002**, *124*, 13185–13193.
23. Griffin, J. R.; Wendell, C. I.; Garwin, J. A.; White, M. C. Catalytic C(sp<sup>3</sup>)–H alkylation via an iron carbene intermediate. *J. Am. Chem. Soc.* **2017**, *139*, 13624–13627.



- 
24. Herschlag, D.; Natarajan, A. Fundamental challenges in mechanistic enzymology: Progress toward understanding the rate enhancements of enzymes. *Biochemistry* **2013**, *52*, 2050–2067.
  25. Prier, C. K.; Zhang, R. K.; Buller, A. R.; Brinkmann-Chen, S.; Arnold, F. H. Enantioselective, intermolecular benzylic C–H amination catalysed by an engineered iron-haem enzyme. *Nat. Chem.* **2017**, *9*, 629–634.
  26. Renata, H.; Lewis, R. D.; Sweredoski, M. J.; Moradian, A.; Hess, S.; Wang, Z. J.; Arnold, F. H. Identification of mechanism-based inactivation in P450-catalyzed cyclopropanation facilitates engineering of improved enzymes. *J. Am. Chem. Soc.* **2016**, *138*, 12527–12533.
  27. Govindaraj, S.; Poulos, T. L. The domain architecture of cytochrome P450BM-3. *J. Biol. Chem.* **1997**, *272*, 7915–7921.
  28. Anslyn, E. V.; Dougherty, D. A. *Modern Physical Organic Chemistry* (University Science, Sausalito, CA, **2006**), chap. 8.
  29. Demonceau, A.; Noels, A. F.; Costa, J.-L.; Hubert, A. J. Rhodium(II) carboxylate-catalyzed reactions of diazoesters: evidence for an equilibrium between free carbene and a metal–carbene complex. *J. Mol. Catal.* **1990**, *58*, 21–26.
  30. Davies, H. M. L.; Hansen, T.; Churchill, M. R. Catalytic asymmetric C–H activation of alkanes and tetrahydrofuran, *J. Am. Chem. Soc.* **2000**, *122*, 3063–3070.
  31. Doyle, M. P.; Westrum, L. J.; Wolthuis, W. N. E.; See, M. M.; Boone, W. P.; Bagheri, V.; Pearson, M. M. Electronic and steric control in carbon-hydrogen insertion reactions of diazoacetoacetates catalyzed by dirhodium(II) carboxylates and carboxamides. *J. Am. Chem. Soc.* **1993**, *115*, 958–964.
  32. Zhang, J.-T.; Qi, X.-L.; Chen, J.; Li, B.-S.; Zhou, Y.-B.; Cao, X.-P. Total synthesis of malyngamides K, L, and 5''-*epi*-C and absolute configuration of malyngamide L. *J. Org. Chem.* **2011**, *76*, 3946–3959.
  33. Edwards, J. T.; Merchant, R. R.; McClymont, K. S.; Knouse, K. W.; Qin, T.; Malins, L. R.; Vokits, B.; Shaw, S. A.; Bao, D.-H.; Wei, F.-L.; Zhou, T.; Eastgate, M. D.; Baran, P. S. Decarboxylative alkenylation. *Nature* **2017**, *545*, 213–218.
  34. He, J.; Hamann, L. G.; Davies, H. M. L.; Beckwith, R. E. J. Late-stage C–H functionalization of complex alkaloids and drug molecules via intermolecular rhodium-carbenoid insertion. *Nat. Commun.* **2015**, *6*, 5943.

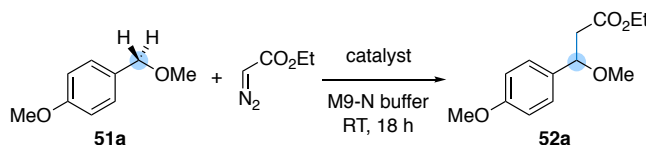
- 
35. Padwa, A.; Hornbuckle, S. F. Ylide formation from the reaction of carbenes and carbenoids with heteroatom lone pairs. *Chem. Rev.* **1991**, *91*, 263–309.
  36. Luo, Y.-R. *Comprehensive Handbook of Chemical Bond Energies* (CRC Press, Boca Raton, FL, **2007**), chap. 3.
  37. Yang, J.-M.; Cai, Y.; Zhu, S.-F.; Zhou, Q.-L. Iron-catalyzed arylation of  $\alpha$ -aryl- $\alpha$ -diazooesters. *Org. Biomol. Chem.* **2016**, *14*, 5516–5519.
  38. Xu, B.; Li, M.-L.; Zuo, X.-D.; Zhu, S.-F.; Zhou, Q.-L. Catalytic asymmetric arylation of  $\alpha$ -aryl- $\alpha$ -diazooacetates with aniline derivatives. *J. Am. Chem. Soc.* **2015**, *137*, 8700–8703.
  39. Cardillo, G.; Gentilucci, L.; Qasem, A. R.; Sgarzi, F.; Spampinato, S. Endomorphin-1 analogues containing  $\beta$ -Proline are  $\mu$ -opioid receptor agonists and display enhanced enzymatic hydrolysis resistance. *J. Med. Chem.* **2002**, *45*, 2571–2578.
  40. Sridharan, V.; Suryavanshi, P. A.; Menéndez, J. C. Advances in the chemistry of tetrahydroquinolines. *Chem. Rev.* **2011**, *111*, 7157–7259.
  41. Kille, S.; Acevedo-Rocha, C. G.; Parra, L. P.; Zhang, Z.-G.; Opperman, D. J.; Reetz, M. T.; Acevedo, J. P. Reducing codon redundancy and screening effort of combinatorial protein libraries created by saturation mutagenesis. *ACS Synth. Biol.* **2013**, *2*, 83–92.
  42. Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchinson III, C. A.; Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **2009**, *6*, 343–345.
  43. Berry, E. A.; Trumpower, B. L. Simultaneous determination of hemes *a*, *b*, and *c* from pyridine hemochrome spectra. *Anal. Biochem.* **1987**, *161*, 1–15.

## SUPPLEMENTARY INFORMATION FOR CHAPTER 4

Material for this chapter appears in **Zhang, R. K.**; Chen, K.; Huang, X.; Wohlschlager, L.; Renata, H.; Arnold, F. H. “Enzymatic assembly of carbon–carbon bonds via iron-catalysed  $sp^3$  C–H functionalisation,” *Nature* **2019**, 565, 67–72. DOI: 10.1038/s41586-018-0808-5. This work was performed in collaboration with all authors. Reprinted with permission from Nature Publishing Group.

## C.1 Supporting Tables C-1 through C-4

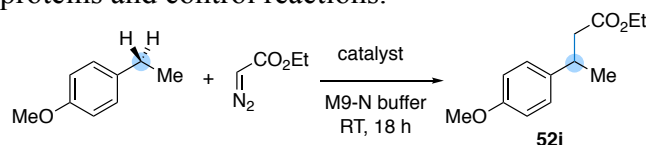
**Table C-1.** Initial results for C–H alkylation of *p*-methoxybenzyl methyl ether (**51a**) with ethyl diazoacetate catalyzed by heme proteins and control reactions.



Catalyst	Catalyst concentration	TTN to <b>52a</b>
P-4 A82L (in <i>E. coli</i> cells, $\text{OD}_{600} = 29$ )	4.0 $\mu\text{M}$	13
<i>R. marinus</i> NOD Y32G (in <i>E. coli</i> cells, $\text{OD}_{600} = 30$ )	11.6 $\mu\text{M}$	7 (12 <sup>†</sup> )
Vector control (in <i>E. coli</i> cells, $\text{OD}_{600} = 40$ ) <sup>*, §</sup>	N. A.	N.D.
Hemin	400 $\mu\text{M}$	N.D. <sup>†, #</sup>
Hemin + BSA	400 $\mu\text{M}$ (hemin), 0.6 mg/mL (BSA)	N.D. <sup>†, #</sup>

**Notes:** Reactions performed with 10 each substrate; results are the average of duplicate reactions. BSA, bovine serum albumin; TTN, total turnover number; N.D., not detected.  $\text{OD}_{600}$  value refers to that of the resuspension used for the reaction. <sup>†</sup>These reactions contain 1 mM  $\text{Na}_2\text{S}_2\text{O}_4$ , used as reductant. <sup>\*</sup>Vector control indicates that *E. coli* harboring pET22b(+) encoding a protein which does not have a transition metal cofactor (halohydrin dehalogenase, UniProt ID: Q93D82) was employed in the reaction. <sup>§</sup>Reaction time was 24 h. <sup>#</sup>For reactions with hemin  $\pm$  BSA, a 0.5 mM solution of hemin ( $\pm$  BSA, 0.75 mg/mL) in M9-N (320  $\mu\text{L}$ ) was used instead of a cell suspension; no *D*-glucose was added to these reactions.

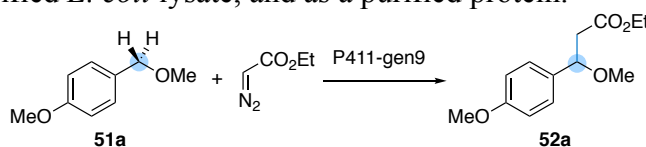
**Table C-2.** Initial results for C–H alkylation of 4-ethylanisole with ethyl diazoacetate catalyzed by heme proteins and control reactions.



Catalyst	Catalyst concentration	TTN to 52i
P-4 A82L (in <i>E. coli</i> cells, OD <sub>600</sub> = 21)	4.0 μM	3
<i>R. marinus</i> NOD Y32G (in <i>E. coli</i> cells, OD <sub>600</sub> = 30)	11.6 μM	< 1 (2 <sup>†</sup> )
Vector control (in <i>E. coli</i> cells, OD <sub>600</sub> = 40)*, ξ	N. A.	N.D.
Hemin	400 μM	N.D. <sup>†, #</sup>
Hemin + BSA	400 μM (hemin), 0.6 mg/mL (BSA)	N.D. <sup>†, #</sup>

**Notes:** Reactions performed with 10 mM each substrate; results are the average of duplicate reactions. Other notes are the same as in Table C-1.

**Table C-3.** P411-gen9, an evolved P411ΔFAD C–H alkylation enzyme, is active in whole *E. coli* cells, in clarified *E. coli* lysate, and as a purified protein.

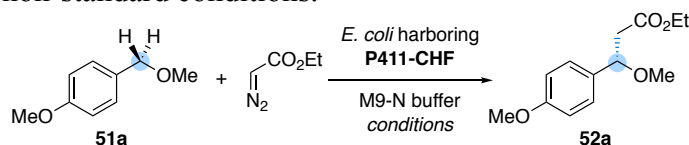


Form	[P411], μM	Exogenous reductant	TTN to 52a
whole <i>E. coli</i> cells (OD <sub>600</sub> = 15–17)	2.0 μM	None	900
Lysate	2.0 μM	Na <sub>2</sub> S <sub>2</sub> O <sub>4</sub> , 1 mM	940
Purified protein	11.9 μM	Na <sub>2</sub> S <sub>2</sub> O <sub>4</sub> , 1 mM	150
Purified protein	11.9 μM	Na <sub>2</sub> S <sub>2</sub> O <sub>4</sub> , 5 mM	210
Purified protein	10.0 μM–11.9 μM	NADPH, 10 mM	250

**Notes:** Reactions were performed with 10 mM each substrate for 18–20 hours at room temperature. Results are the average of at least two reactions.

These experiments demonstrate that C–H alkylation can be catalyzed using an evolved P411ΔFAD enzyme in whole *E. coli* cells, in clarified *E. coli* lysate, and as a purified protein. The results using purified P411-gen9 are preliminary and the TTN values should be interpreted with caution as no studies regarding the effect of purification conditions on the activity of the enzyme nor studies to optimize the reaction conditions were pursued.

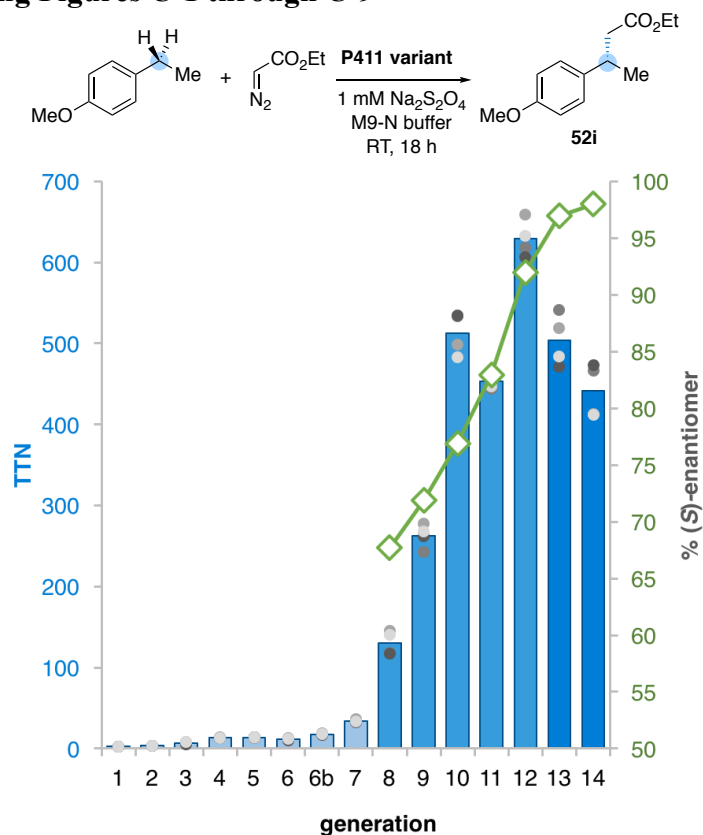
**Table C-4.** Enzymatic C–H alkylation reactions performed using *E. coli* cells harboring P411-CHF under non-standard conditions.



Conditions	TTN to 52a	e.r.
Anaerobic, full system, room temp.	2150	96.7 : 3.3
Anaerobic, full system, 4 °C	2090	98.0 : 2.0
Anaerobic, no GOX, room temp.	2100	96.7 : 3.3
Anaerobic, no GOX, no <i>D</i> -glucose, room temp.	1770	96.7 : 3.3
Aerobic, no GOX, no <i>D</i> -glucose, room temp.	30	not measured

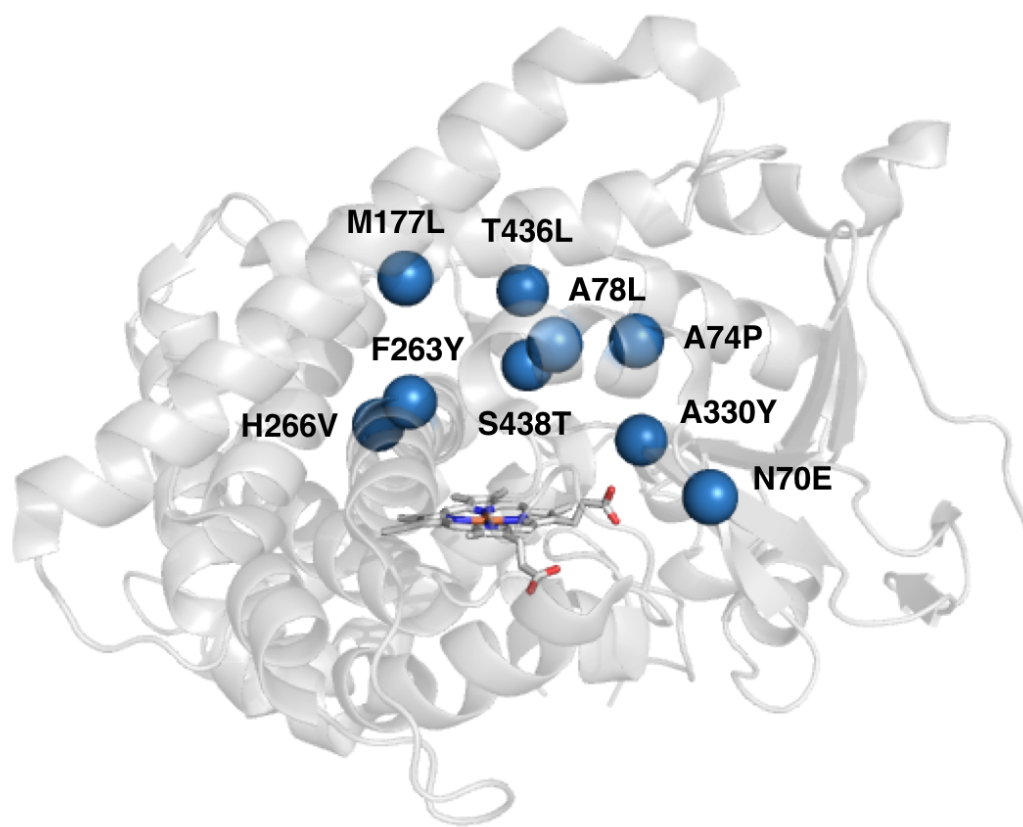
**Notes:** Standard reaction conditions are *E. coli* cells expressing P411-CHF (resuspended to OD<sub>600</sub> = 30) with 10 mM each substrate. Reactions were allowed to proceed at the indicated temperature for 18 h. For reactions without the GOX oxygen depletion solution, 20 µL of M9-N buffer was added instead; for reactions without *D*-glucose, 40 µL of M9-N buffer was added. The aerobic reaction was set up under air, without prior degassing of cells or reagents with argon. Results are the average of duplicate reactions.

## C.2 Supporting Figures C-1 through C-9



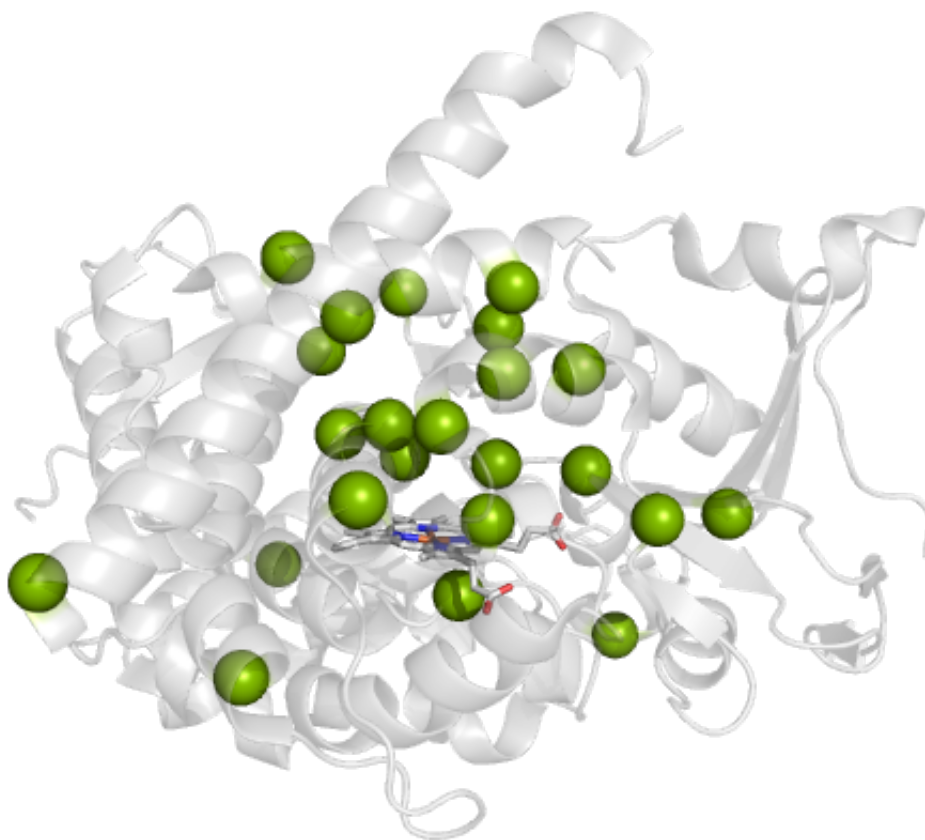
**Figure C-1.** Evolutionary lineage from P-4 A82L to P411-CHF evaluated for C–H alkylation of 4-ethylanisole.

**Notes:** Standard reaction conditions: clarified lysate of *E. coli* expressing the indicated heme protein variant, 10 mM of each substrate, 1 mM Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub>. Reactions with generations 1, 2, and 3 variants employed 4.0 μM heme protein; all other reactions used 2.0 μM heme protein. Bars represent mean TTN values averaged over four reactions (performed from two independent cell cultures, each used for duplicate reactions); each TTN data point is shown as a grey dot. Enantioselectivity results are represented by green diamonds. See Table C-12 in **Section C.III** for the data presented here. TTN, total turnover number; RT, room temperature.



**Figure C-2.** Structural visualization of amino acids mutated during directed evolution of P-4 A82L to P411-CHF.

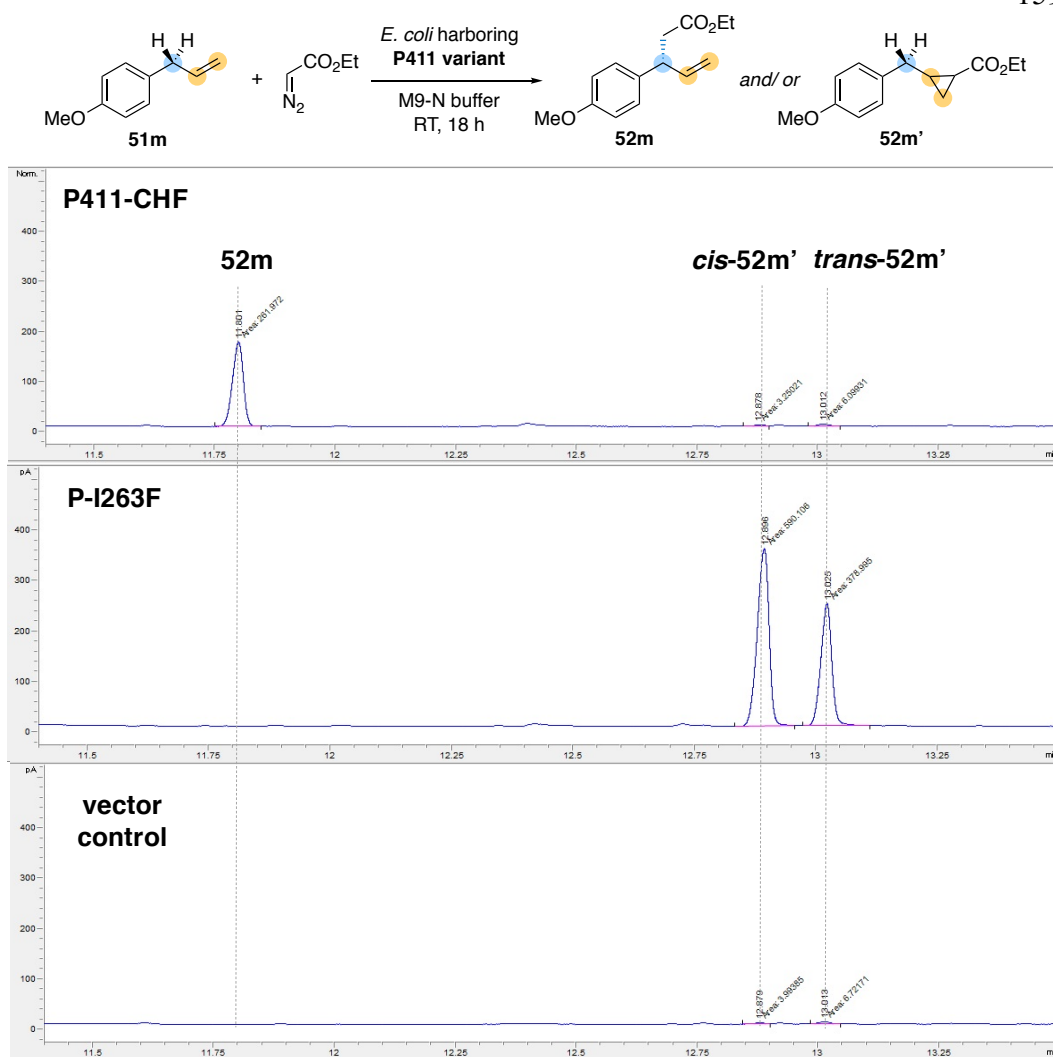
**Notes:** The structure of P-4 A82L (heme domain) was modeled using the crystal structure of a related P411 variant (PDB 5UCW), which contains two additional mutations. Considering only the changes incurred in the heme domain, the following mutations were accumulated in going from P-4 A82L to P411-CHF: N70E, A74P, A78L, M177L, F263Y, H266V, A330Y, T436L, S438T (shown as blue spheres, residues 327 and 437 were not included in this analysis because P-4 A82L and P411-CHF contain the same amino acid residues at those positions). Most of the mutations are at positions that line the distal heme pocket and all of the mutated residues are within 15 Å of the iron atom in the heme cofactor.



**Figure C-3.** Structural visualization of amino acid differences between P450<sub>BM3</sub> wild-type and P411-CHF.

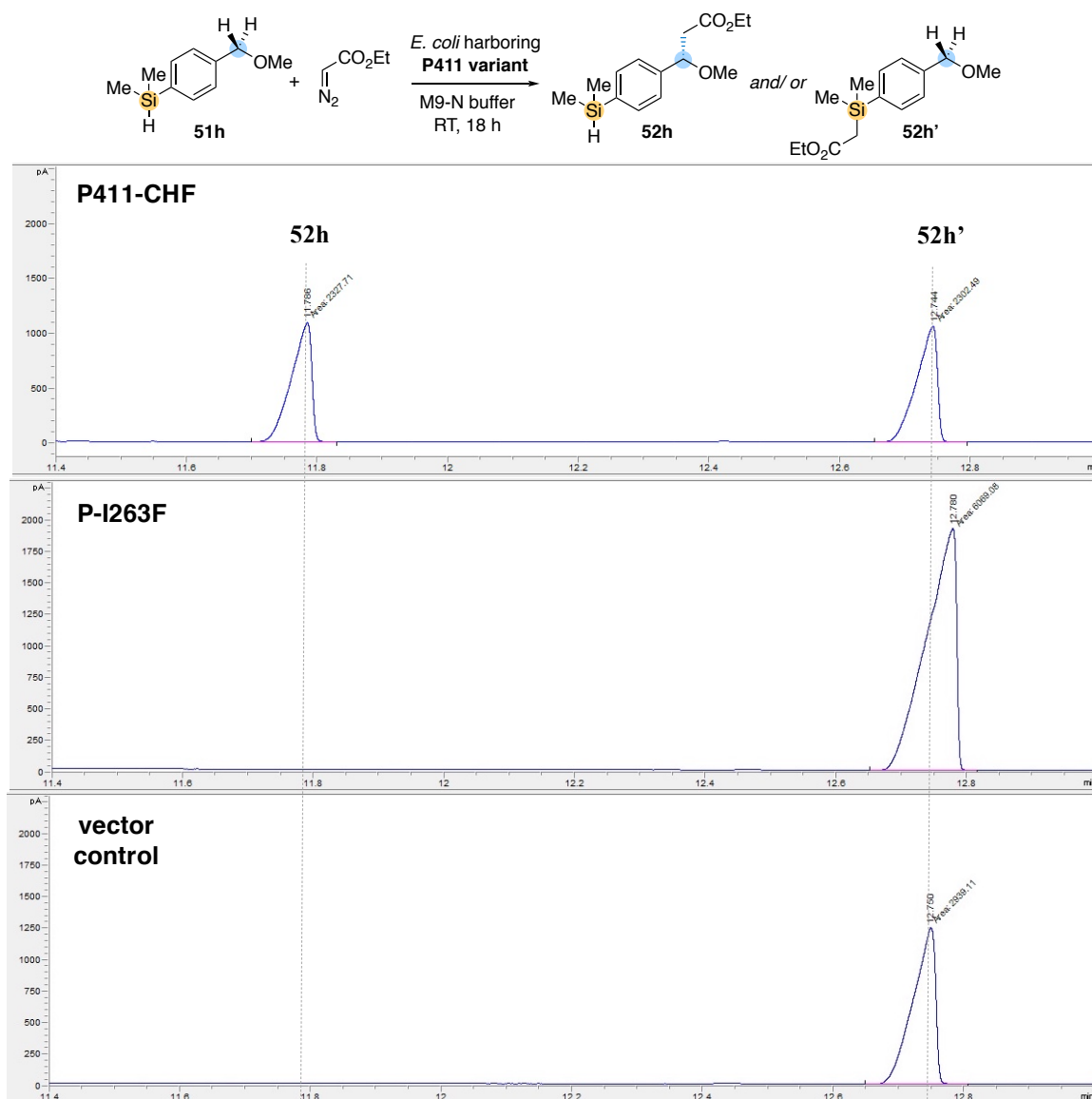
**Notes:** The structure of P450<sub>BM3</sub> wild-type (heme domain) was modeled using the crystal structure of H. M. Girvan *et al.* (PDB 2IJ2). Comparing only the heme domains of the two proteins, the following 23 amino acids are changed in P411-CHF relative to P450<sub>BM3</sub> wild-type: N70E, A74P, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, A328V, A330Y, L353V, I366V, C400S, T436L, E442K (shown as green spheres). Overall, 5% of the amino acids in the heme domain of P450<sub>BM3</sub> have been substituted in P411-CHF.





**Figure C-4.** Product profiles for the reaction of 4-allylanisole (**51m**) and ethyl diazoacetate with P411-CHF and with P-I263F.

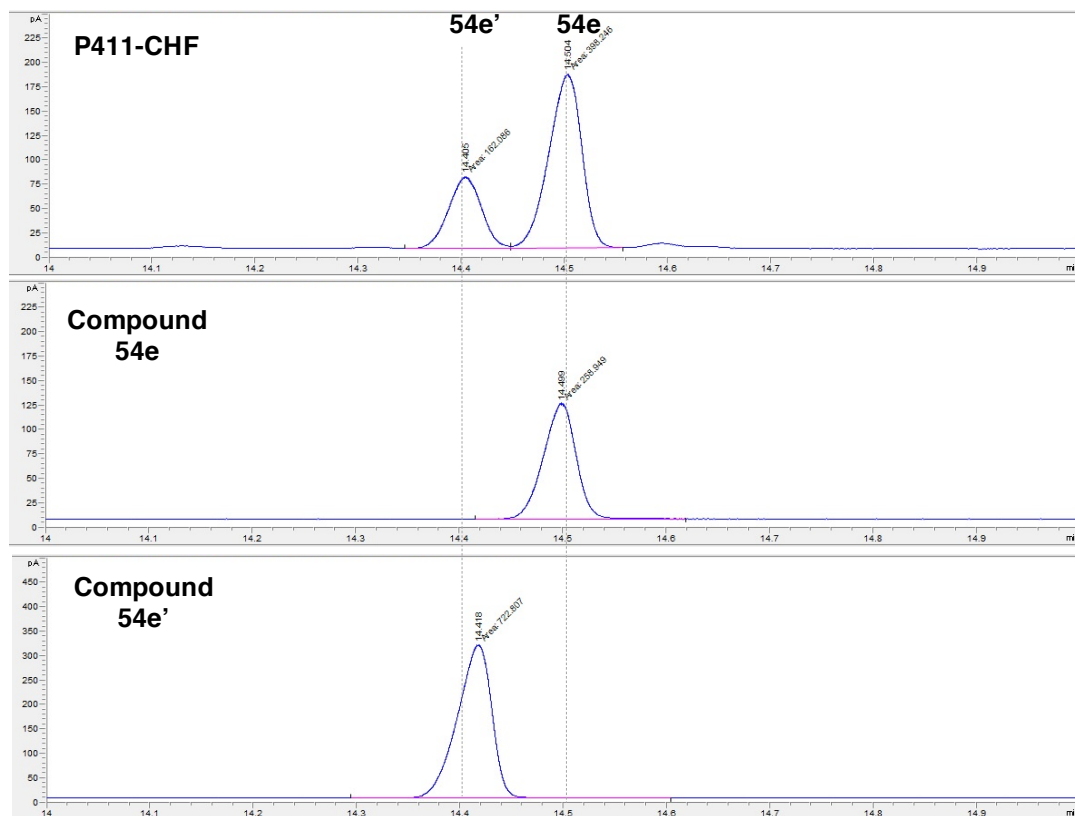
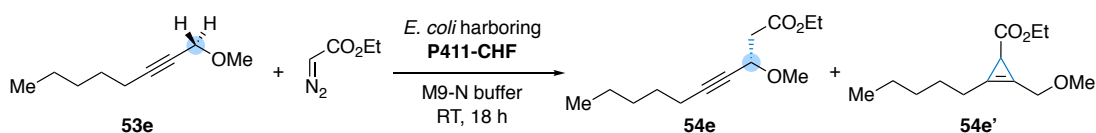
**Notes:** Representative GC traces of reactions performed with P411-CHF (top), P-I263F (middle), or a vector control (bottom) in *E. coli* cells. Reactions were performed with 10 mM each substrate. Vector control indicates that *E. coli* harboring pET22b(+) encoding a protein that does not contain a transition metal cofactor (halohydrin dehalogenase, UniProt ID: Q93D82) was employed in the reaction.



**Figure C-5.** P411-CHF catalyzes C–H alkylation of **51h** with ethyl diazoacetate.

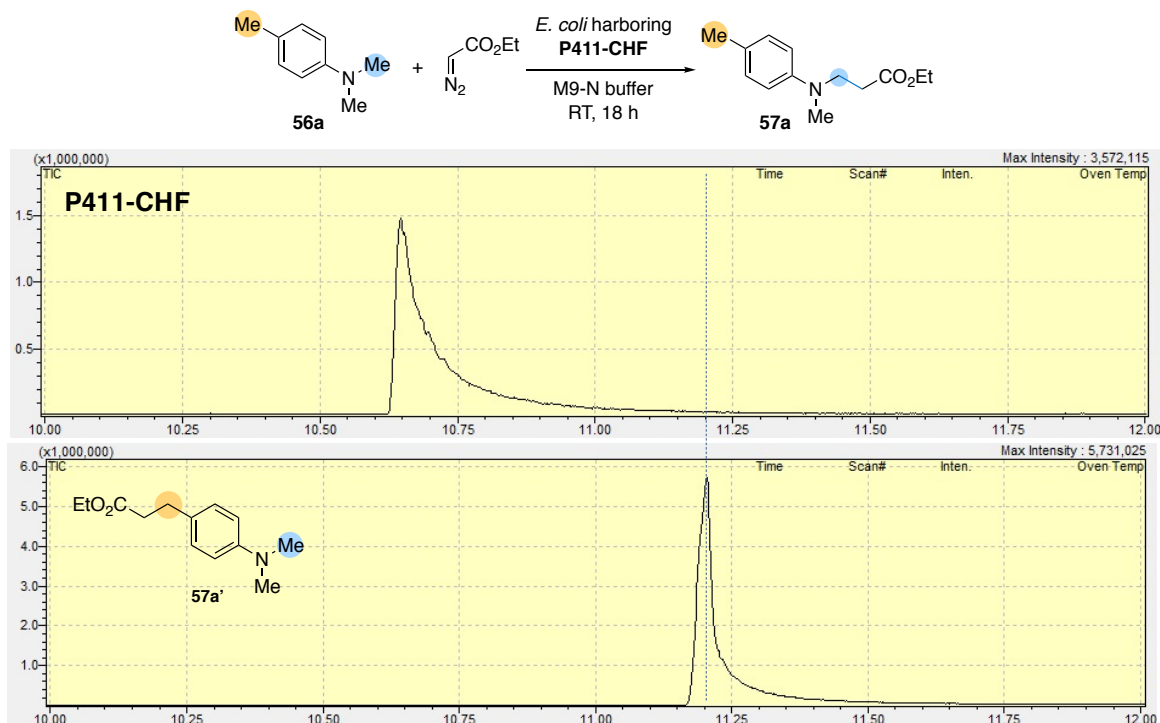
**Notes:** Representative GC traces of reactions performed with P411-CHF (top), P-I263F (middle), or a vector control (bottom) in *E. coli* cells. Reactions were performed as in **Supplementary Fig. 6**.

While P411-CHF was able to access C–H alkylation product **52h** (930 TTN to **52h**, 1:1 **52h**:**52h'**) in the presence of a reactive Si–H bond, reaction with *E. coli* harboring P-I263F only afforded Si–H insertion product **52h'**. Additionally, compound **52h'** was also observed in vector control reactions employing whole *E. coli* cells. Neither **52h** nor **52h'** were detected in a control reaction where **51h** and ethyl diazoacetate were combined in M9-N buffer in the absence of cells.



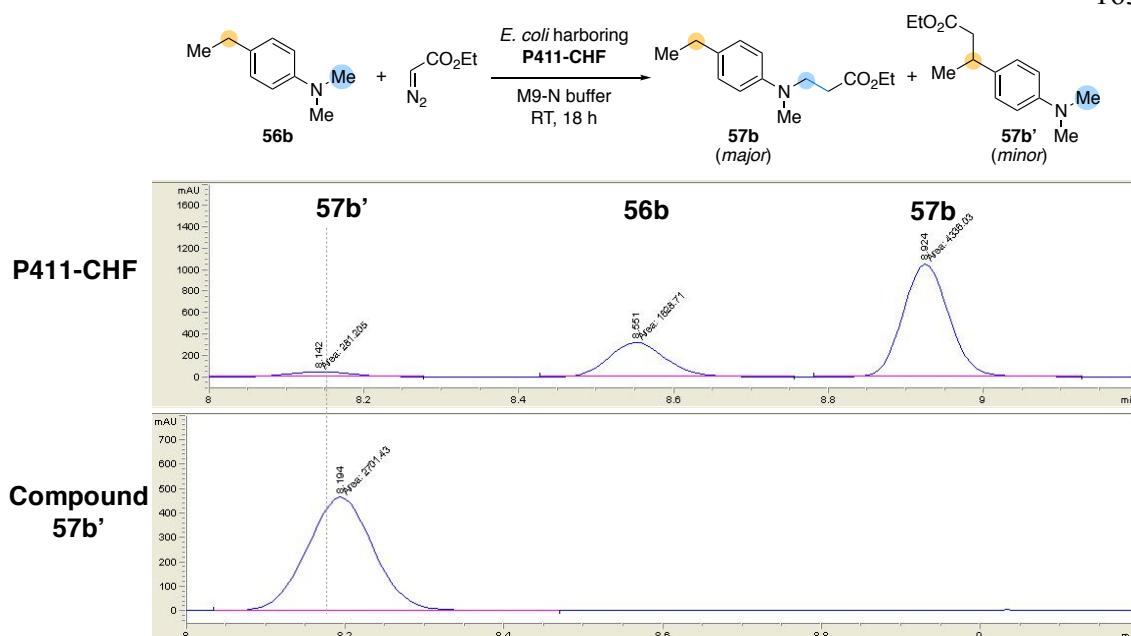
**Figure C-6.** Enzymatic reaction of P411-CHF with substrate **53e** and ethyl diazoacetate produces cyclopropene **54e'** as a side product.

**Notes:** Representative GC trace of a reaction performed using P411-CHF in *E. coli* cells (top) with 10 mM each substrate. Also shown are GC traces for reference compounds **54e** (middle) and **54e'** (bottom), used for the identification of P411-CHF synthesized products.



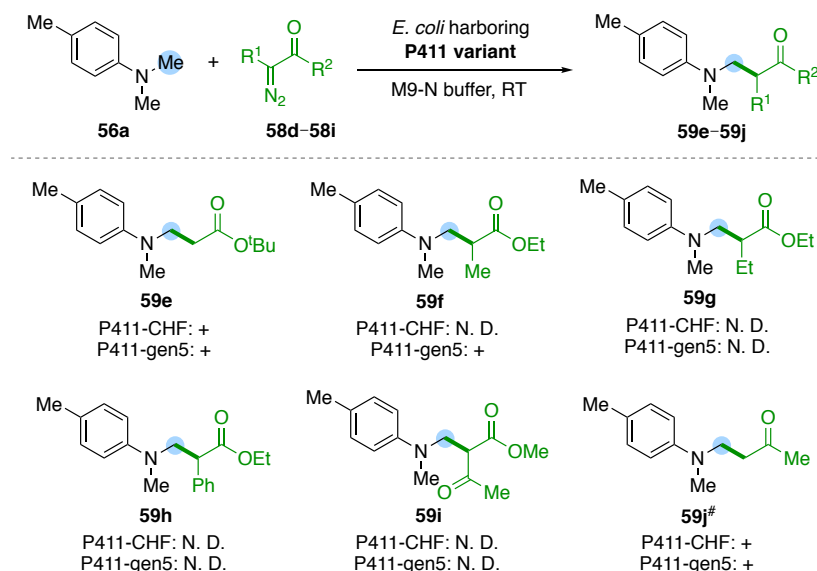
**Figure C-7.** Enzymatic C–H alkylation of 4,*N,N*-trimethylaniline (**56a**) with P411-CHF is selective for the  $\alpha$ -amino position.

**Notes:** Representative GC-MS trace of a reaction performed using P411-CHF in *E. coli* cells (top) with 10 mM each substrate. Also shown is a trace for reference compound **57a'** (bottom), which is not observed in the reaction with P411-CHF. This demonstrates that P411-CHF has exquisite selectivity for the alkylation of  $\alpha$ -amino C–H bond in the presence of primary benzylic C–H bonds.



**Figure C-8.** Product profile of P411-CHF catalyzed C–H alkylation of 4-ethyl-*N,N*-dimethylaniline (**56b**) with ethyl diazoacetate.

**Notes:** Representative HPLC trace of a reaction performed using P411-CHF in *E. coli* cells (top) with 10 mM each substrate. Also shown is a trace for reference compound **57b'** (bottom), which was identified as a minor side product in the P411-CHF catalyzed reaction. Major product **57b** was assigned by performing a preparative scale enzymatic reaction (see Section C.8) and isolation and characterization of the major product.



**Figure C-9.** Additional diazo substrates tested for C–H alkylation with P411-CHF and P411-gen5. ‘+’ indicates product was detected; N. D., not detected. <sup>#</sup>Other products derived from compound **56a** were also observed by GC-MS.

**Notes:** Reactions were performed using *E. coli* cells expressing P411-CHF (resuspended to OD<sub>600</sub> = 59) or P411-gen5 (resuspended to OD<sub>600</sub> = 49) with 10 mM coupling partner **56a** and 10 mM diazo compound. Product formation was analyzed by GC-MS only. The identity of the product was not confirmed by comparison with chemically synthesized reference compounds or through isolation and characterization. These preliminary results are noteworthy, but should not be used alone for drawing conclusions.

Diazo compounds **58a**, **58b**, and **58c**, for the formation of products **59a**, **59b**, and **59c**, were also investigated with P411-CHF (+, +, +, respective preliminary results) and P411-gen5 (+, N. D., +). Additional variants were tested for these transformations and subsequent preparative scale reactions for product isolation and characterization were pursued; these results are described in Figure 4-8.

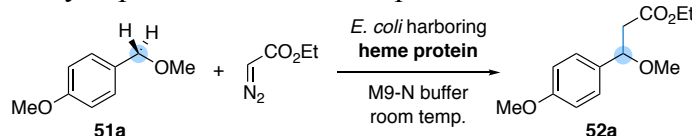
Because P411-CHF and its related variants are fully genetically encoded and can be readily optimized by directed evolution, these enzymes can serve as starting points for the evolution of future variants. Some of these future variants will certainly surpass P411-CHF in C–H alkylation activity with alternative diazo reagents.

### C.3 Screening of heme proteins for C–H alkylation activity

#### C.3.1 Testing diverse heme proteins for reaction discovery

A composite plate of 40 heme proteins and their variants from various organisms were screened for formation of product **52a** from substrates **51a** and ethyl diazoacetate. These proteins were cloned and used in other studies, including carbene Si–H insertion<sup>1</sup> and alkene cyclopropanation<sup>2,3</sup>. Expression of these proteins followed the procedures as described by the prior studies and testing for initial activity was carried out with whole *E. coli* cells. The general procedure for reaction screening in 96-well plate format was employed and the reactions were analyzed by GC-MS. *Note:* Since this was an initial test, proper expression of the indicated proteins was not verified. Consequently, negative results should be interpreted with caution. From this experiment, it was observed that *R. marinus* NOD (Y32G) had a low level of C–H alkylation activity (Table C-5).

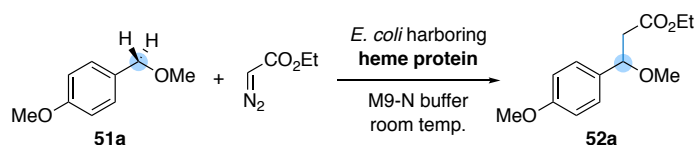
In addition, eight diverse heme proteins, including *R. marinus* NOD (Y32G), were chosen and tested for C–H alkylation activity using substrates **51a** and ethyl diazoacetate following the general procedure for small scale enzymatic reactions using whole *E. coli* cells (Table C-6). Product formation by variant *R. marinus* NOD (Y32G) was characterized by HPLC. Other reactions were analyzed by GC-MS for formation of product **52a** and the expression of all proteins were measured using the hemochrome assay. Conditions with no exogenous reductant as well as with the addition of 1 mM Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub> (final concentration) were tested.

**Table C-5.** Preliminary experiments with heme proteins.

UniProt ID	Organism	Annotation	Abbrev. name	No. of variants	Ref.	Formation of 3a
B3FQS5	<i>Rhodothermus marinus</i>	Cytochrome <i>c</i>	<i>Rma</i> cyt <i>c</i>	9	1	n. d.
P15452	<i>Hydrogenobacter thermophilus</i>	Cytochrome <i>c</i> -552	<i>Hth</i> cyt <i>c</i>	2	1	n. d.
P00080	<i>Rhodopila globiformis</i>	Cytochrome <i>c</i> 2	-	2	1	n. d.
D0MGT2	<i>Rhodothermus marinus</i>	Nitric oxide dioxygenase	<i>Rma</i> NOD	10	11	detected <sup>†</sup>
P02185	<i>Physeter catodon</i> (Sperm whale)	Myoglobin	Mb	1	2*	n. d.
B3DUZ7	<i>Methylobacterium inferorum</i>	Hemoglobin-like flavoprotein	HGG <sup>#</sup>	2	2*	n. d.
O66586	<i>Aquifex aeolicus</i>	Thermoglobin	-	1	2	n. d.
G7VHJ7	<i>Pyrobaculum ferrireducens</i>	Protoglobin	-	1	2	n. d.
Q0PB48	<i>Campylobacter jejuni</i>	Truncated hemoglobin	-	1	2	n. d.
Q9YFF4	<i>Aeropyrum pernix</i>	Protoglobin	-	10	2	n. d.
O31607	<i>Bacillus subtilis</i>	Truncated hemoglobin	-	1	3	n. d.

**Notes:** Reactions performed with 10 mM each substrate. No., number; Abbrev., abbreviated; Ref., reference.; n. d., not detected. <sup>†</sup>Small amount of **52a** was detected in the reaction with variant *R. marinus* NOD (Y32G). No product **52a** was detected in the reaction mixtures of other *R. marinus* NOD variants. \*While these heme proteins were not reported in this reference, the expression conditions employed for these variants followed those given in the reference. <sup>#</sup>This protein is abbreviated as Hell's Gate Globin (HGG) on the basis of a prior literature report which uses this name<sup>4</sup>.



**Table C-6.** Heme proteins tested for C–H alkylation activity.

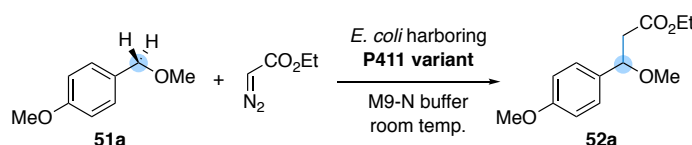
UniProt ID	Organism	Annotation	Abbrev. Name	Mutation(s) from WT	Formation of 3a
P14779	<i>Bacillus megaterium</i>	cytochrome P450/NADPH-P450 reductase	P450 <sub>BM3</sub>	none <sup>*</sup>	n. d.
Q55080	<i>Sulfolobus acidocaldarius</i>	Cytochrome P450	CYP119	none <sup>*</sup>	n. d.
D0MGT2	<i>Rhodothermus marinus</i>	Nitric oxide dioxygenase	<i>Rma</i> NOD	Y32G <sup>+, *</sup>	7 TTN <sup>†</sup> (12 TTN <sup>ξ</sup> )
B3DUZ7	<i>Methylobacterium thermophilum</i>	Hemoglobin-like flavoprotein	HGG	none <sup>+, *</sup>	n. d.
P02185	<i>Physeter catodon</i> (Sperm whale)	Myoglobin	Mb	H64V V68A D122N <sup>+, *</sup>	n. d.
B3FQS5	<i>Rhodothermus marinus</i>	Cytochrome <i>c</i>	<i>Rma</i> cyt <i>c</i>	none <sup>+, #</sup>	n. d.
B3FQS5	<i>Rhodothermus marinus</i>	Cytochrome <i>c</i>	<i>Rma</i> cyt <i>c</i>	V75T M100D M103E <sup>+, #</sup>	n. d.
P15452	<i>Hydrogenobacter thermophilus</i>	Cytochrome <i>c</i>	<i>Hth</i> cyt <i>c</i>	none <sup>+, #</sup>	n. d.

**Notes:** Heme proteins were tested in whole *E. coli* cells (OD<sub>600</sub> = 30) both with and without the addition of 1 mM Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub> (final concentration). Reactions were performed with 10 mM each substrate; results are the average of at least duplicate reactions. Abbrev., abbreviated; WT, wild-type; n. d., not detected; TTN, total turnover number. <sup>†</sup>This data is also included in Table C-1. <sup>ξ</sup>Reaction includes 1 mM Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub>. <sup>+</sup>These variants are also included in Table C-5. <sup>\*</sup>These proteins were expressed following the procedure of A. M. Knight *et al.*<sup>2</sup>. <sup>#</sup>Cytochrome *c* variants were expressed following the procedure of S. B. J. Kan *et al.*<sup>1</sup>

### C.3.2 Screening previously engineered cytochrome P450<sub>BM3</sub> variants

A composite plate of 36 cytochrome P450<sub>BM3</sub> variants from lineages engineered for non-natural reactions including C–H amination<sup>5</sup>, olefin aziridination<sup>6</sup>, and olefin cyclopropanation<sup>7a</sup> was screened for C–H alkylation activity using substrates **51a** and ethyl diazoacetate. Variants of cytochrome P450<sub>BM3</sub> in which the axial cysteine ligand is replaced with serine are termed cytochrome P411<sup>16b</sup>. The general procedure for reaction screening in 96-well plate format was employed and the reactions were analyzed by HPLC. Most variants showed no activity (17 variants) or trace activity (15 variants) for **52a** formation. The four highest performing variants, all which contain an axial cysteine-to-serine mutation (P411), are shown in Table C-7. While P-4 A82L was chosen as the parent protein for the directed evolution of a C–H alkylation enzyme, the information gained from this experiment was used to guide the first several rounds of evolution.

**Table C-7.** Engineered cytochrome P411 variants show promiscuous C–H alkylation activity.



Name	Mutations from P450 <sub>BM3</sub> WT	Relative amount of <b>52a</b> formed <sup>†</sup>
P-4 (A82L)	V78A, A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263F, T268G, A290V, A328V, L353V, I366V, C400S, T438S, E442K	1
P-4 (A82L A78V F263Y)	A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263Y, T268G, A290V, A328V, L353V, I366V, C400S, T438S, E442K	1.4
P-4 (A82L A78V F263L)	A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263L, T268G, A290V, A328V, L353V, I366V, C400S, T438S, E442K	0.5
P-4 (A82L A78V F263L E267D) (a.k.a. P411 <sub>CHA</sub> )	A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263L, E267D, T268G, A290V, A328V, L353V, I366V, C400S, T438S, E442K	0.7

**Notes:** Reactions were performed with 10 mM each substrate following the general procedure for reaction screening in 96-well plate format. Results are the average of duplicate reactions. All variants are reported in ref. 5. <sup>†</sup>This refers to the ratio of the HPLC peak area for **52a** in reactions with the indicated variant to that in reactions with P-4 A82L.

#### C.4 Directed evolution of C–H alkylation enzymes

Site-saturation libraries were generated employing the “22c-trick” method<sup>8</sup> and screened in one 96-well plate; double site-saturation libraries were generated using the same method to target two different sites and these were screened in three 96-well plates. Following the general screening in 96-well plate procedure, variants which exhibited higher formation of C–H alkylation product (**52a** or **52i**) or improved enantioselectivity for product **3a** were identified. A summary of the amino acid residues targeted for mutagenesis is presented in Table C-8, as well as the beneficial mutation(s) selected for each round of mutagenesis. The locations of the selected beneficial mutations are displayed on a structural model of the P411 enzyme (Section C.2, Table C-2).

Variants which were identified to show higher activity and/ or enantioselectivity during screening were streaked out on LB<sub>amp</sub> agar plates. A single colony was selected, sequenced, and the TTN measured for both products **52a** and **52i** using clarified lysate of *E. coli* cells overexpressing the desired protein (unless otherwise indicated, the concentration of P411 variant was normalized such that each reaction contained 2.0  $\mu$ M enzyme. Enantiomeric ratios of the enzymatic products produced by P411-gen6 and further evolved variants were also characterized. The results are summarized in Tables C-11 and C-12.

**Table C-8.** Summary of directed evolution for C–H alkylation.

Round	Parent	Diversification strategy <sup>†</sup>	Screening substrate (selection criteria)	Changes made <sup>*</sup>
1	<b>P-4 A82L</b>	Individual variants identified as active for C–H amination <sup>14</sup>	4-ethylanisole (activity)	F263Y
2	<b>P-4 A82L F263Y</b>	Site-saturation mutagenesis A78X	4-ethylanisole (activity)	A78L
3	<b>P-4 A82L F263Y A78L</b>	Site-saturation mutagenesis T327X	4-ethylanisole (activity)	T327I
4	<b>P411-gen4</b> (P-4 A82L F263Y A78L T327I)	Site-saturation mutagenesis A74X, E267X	4-ethylanisole (activity)	A74G
5	<b>P411-gen5</b> (P411-gen4 A74G)	Site-saturation mutagenesis A328X, H92X, R255X, A264X, H100X, F393X, L437X	<i>p</i> -methoxybenzyl methyl ether (activity)	L437Q
6	<b>P411-gen6</b> (P411-gen5 L437Q)	Protein truncations: full-length P411, ΔFAD domain, heme-domain only	<i>p</i> -methoxybenzyl methyl ether (activity)	ΔFAD domain
6b	<b>P411-gen6b</b> (P411ΔFAD-gen6)	Site-saturation mutagenesis A78X, F87X, I263X, T438X	<i>p</i> -methoxybenzyl methyl ether (activity)	S438T
7	<b>P411-gen7</b> (P411-gen6b S438T)	Site-saturation mutagenesis <sup>ψ</sup> A330X, F331X, T436X, A82X, L181X, L188X	<i>p</i> -methoxybenzyl methyl ether (activity)	T436L
8	<b>P411-gen8</b> (P411-gen7 T436L)	Site-saturation mutagenesis D63X, F162X, M177X, V178X, L439X	4-ethylanisole (activity)	M177L
9	<b>P411-gen9</b> (P411-gen8 M177L)	Site-saturation mutagenesis F87X, E267X, M118X, L437X, H266X, S332X, T260X, T365X	4-ethylanisole (activity)	H266V
10	<b>P411-gen10</b> (P411-gen9)	Site-saturation mutagenesis N70X, T88X, H171X,	4-ethylanisole (activity)	N70E

	H266V)	H361X, P329X, T269X, L75X, L52X	& <i>p</i> -methoxybenzyl methyl ether (enantioselectivity)	
11	<b>P411-gen11</b> (P411-gen10 N70E)	Site-saturation mutagenesis L71X, S72X, F261X, G265X, L86X, I401X, A330X, C400X	<i>p</i> -methoxybenzyl methyl ether (activity & enantioselectivity)	A330Y
12	<b>P411-gen12</b> (P411-gen11 A330Y)	Double site-saturation mutagenesis P329X-F331X <sup>#</sup> , A328X- F331X <sup>#</sup> , T327X-T268X, A74X-L437X	<i>p</i> -methoxybenzyl methyl ether (activity & enantioselectivity)	I327T
13a	<b>P411-gen13</b> (P411-gen12 I327T)	Double site-saturation mutagenesis L181X-L437X, V178X- E267X, M118X-I401X	<i>p</i> -methoxybenzyl methyl ether (activity & enantioselectivity)	None
13b	<b>P411-gen13</b> (P411-gen12 I327T)	Testing previously identified beneficial and neutral mutations <sup>ξ</sup> .	<i>p</i> -methoxybenzyl methyl ether (activity & enantioselectivity)	G74P, Q437L
14	<b>P411-CHF</b> (P411-gen13 G74P Q437L)	N. A.	N. A.	N. A.

**Notes:** Some residues were saturated more than once, in different parents. Gen, generation; N. A., not applicable. <sup>†</sup>Residues for site-saturation mutagenesis libraries are listed relative to the amino acid at that position in wild-type P450<sub>BM3</sub>. \*Beneficial mutations are listed relative to the amino acid at that position in the parent protein. <sup>ψ</sup>Random mutagenesis by error-prone PCR on this parent enzyme and screening for C–H alkylation activity was also performed (unpublished results). However, no mutations from this study were carried forward to the next enzyme generation; the F162L mutation, identified in this context, was included in the diversification strategy of round 13b. <sup>#</sup>Only NDT libraries were constructed and screened for this double-site saturation experiment. <sup>ξ</sup>Select mutations identified from previous rounds of mutagenesis and screening were introduced in various combinations to P411-gen13. Twenty-seven variants were attempted and screened.

**Table C-9.** Mutations and truncations relative to P450<sub>BM3</sub> wild-type for the P411 variants described in Table C-8.

Gen.	Name	Mutations relative to P450 <sub>BM3</sub> WT	Domain composition <sup>†</sup>
1	P-4 A82L	V78A, A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263F, T268G, A290V, A328V, L353V, I366V, C400S, T438S, E442K	Full-length
2	P-4 A82L F263Y	V78A, A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263Y, T268G, A290V, A328V, L353V, I366V, C400S, T438S, E442K	Full-length
3	P-4 A82L F263Y A78L	V78L, A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263Y, T268G, A290V, A328V, L353V, I366V, C400S, T438S, E442K	Full-length
4	<b>P411-gen4</b> (P-4 A82L F263Y A78L T327I)	V78L, A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263Y, T268G, A290V, T327I, A328V, L353V, I366V, C400S, T438S, E442K	Full-length
5	<b>P411-gen5</b> (P411-gen4 A74G)	A74G, V78L, A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263Y, T268G, A290V, T327I, A328V, L353V, I366V, C400S, T438S, E442K	Full-length
6	<b>P411-gen6</b> (P411-gen5 L437Q)	A74G, V78L, A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263Y, T268G, A290V, T327I, A328V, L353V, I366V, C400S, L437Q, T438S, E442K	Full-length
6b	<b>P411-gen6b</b> (P411ΔFAD–gen6)	A74G, V78L, A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263Y, T268G, A290V, T327I, A328V, L353V, I366V, C400S, L437Q, T438S, E442K	ΔFAD domain
7	<b>P411-gen7</b> (P411-gen6b S438T)	A74G, V78L, A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263Y, T268G, A290V, T327I, A328V, L353V, I366V, C400S, L437Q, E442K	ΔFAD domain
8	<b>P411-gen8</b> (P411-gen7 T436L)	A74G, V78L, A82L, F87A, P142S, T175I, A184V, S226R, H236Q, E252G, I263Y, T268G, A290V, T327I, A328V, L353V, I366V, C400S, T436L, L437Q, E442K	ΔFAD domain
9	<b>P411-gen9</b> (P411-gen8)	A74G, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G,	ΔFAD domain

	M177L)	I263Y, T268G, A290V, T327I, A328V, L353V, I366V, C400S, T436L, L437Q, E442K	
10	<b>P411-gen10</b> (P411-gen9 H266V)	A74G, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, T327I, A328V, L353V, I366V, C400S, T436L, L437Q, E442K	ΔFAD domain
11	<b>P411-gen11</b> (P411-gen10 N70E)	N70E, A74G, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, T327I, A328V, L353V, I366V, C400S, T436L, L437Q, E442K	ΔFAD domain
12	<b>P411-gen12</b> (P411-gen11 A330Y)	N70E, A74G, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, T327I, A328V, A330Y, L353V, I366V, C400S, T436L, L437Q, E442K	ΔFAD domain
13	<b>P411-gen13</b> (P411-gen12 I327T)	N70E, A74G, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, A328V, A330Y, L353V, I366V, C400S, T436L, L437Q, E442K	ΔFAD domain
14	<b>P411-CHF</b> (P411-gen13 G74P Q437L)	N70E, A74P, V78L, A82L, F87A, P142S, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, A328V, A330Y, L353V, I366V, C400S, T436L, E442K	ΔFAD domain

<sup>†</sup>Full-length cytochrome P411 variants contain amino acids 1–1048. Cytochrome P411 variants containing the FAD truncation (ΔFAD domain) contain amino acids 1–664.

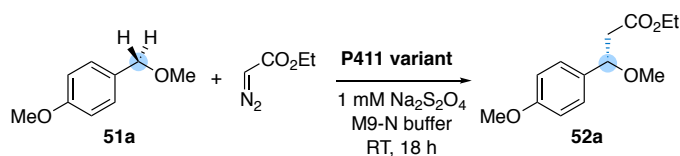
**Table C-10.** Other P411 variants described in this study.

Name	Mutations from P450 <sub>BM3</sub> WT	Reference
P-I263F	V78A, F87V, P142S, T175I, A184V, S226R, H236Q, E252G, I263F, T268A, A290V, L353V, I366V, C400S, T438S, E442K	9
P411-IY T327I	N70E, A74G, V78L, A82L, F87A, M118S, P142S, F162L, T175I, M177L, A184V, S226R, H236Q, E252G, I263Y, H266V, T268G, A290V, T327I, A328V, A330Y, L353V, I366V, C400S, I401L, T436L, L437Q, E442K	N. A. <sup>†</sup>

<sup>†</sup>N. A., not applicable. The parent of this variant, P411-IY, was made for round 13b of

directed evolution (see Table C-8). The P411-IY T327I variant, which contains an additional T327I mutation, was identified in the context of a separate carbene transfer project (Chen, K., Zhou, A. Z. & Arnold, F. H., unpublished results).

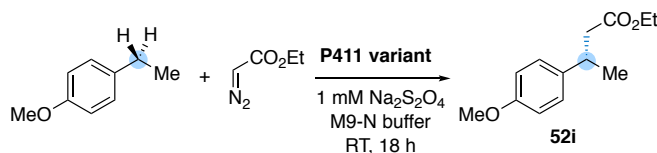
**Table C-11.** Enzymatic C–H alkylation data presented in Figure 4-4.



Variant	[P411], $\mu\text{M}$	TTN $\pm$ SD	e.r.
P-4 A82L	4.0	14 $\pm$ 2	N. A.
P-4 A82L F263Y	4.0	29 $\pm$ 1	N. A.
P-4 A82L F263Y A78L	4.0	23 $\pm$ 5	N. A.
P411-gen4	2.0	48 $\pm$ 3	N. A.
P411-gen5	2.0	68 $\pm$ 3	N. A.
P411-gen6	2.0	59 $\pm$ 4	<i>rac</i>
P411-gen6b	2.0	98 $\pm$ 2	<i>rac</i>
P411-gen7	2.0	200 $\pm$ 4	55.0 : 45.0
P411-gen8	2.0	560 $\pm$ 50	59.4 : 40.6
P411-gen9	2.0	940 $\pm$ 30	64.7 : 35.3
P411-gen10	2.0	1480 $\pm$ 50	68.1 : 31.9
P411-gen11	2.0	1240 $\pm$ 30	77.9 : 22.1
P411-gen12	2.0	1490 $\pm$ 40	88.5 : 11.5
P411-gen13	2.0	1960 $\pm$ 40	94.0 : 6.0
P411-CHF	2.0	2020 $\pm$ 40	96.7 : 3.3

**Notes:** Reactions were performed using clarified *E. coli* lysate with 10 mM each substrate. Each reported TTN is the average over four reactions (performed from two independent cell cultures, each used for duplicate reactions). TTN, total turnover number, refers to TTN to **52a**; SD, standard deviation; N. A., not applicable, indicates that this value was not measured; *rac*, racemic.



**Table C-12.** Enzymatic C–H alkylation data presented in Figure C-1.

Variant	[P411], $\mu\text{M}$	TTN $\pm$ SD	e.r.
P-4 A82L	4.0	$2 \pm 0$	N. A.
P-4 A82L F263Y	4.0	$4 \pm 0$	N. A.
P-4 A82L F263Y A78L	4.0	$7 \pm 2$	N. A.
P411-gen4	2.0	$13 \pm 1$	N. A.
P411-gen5	2.0	$14 \pm 0$	N. A.
P411-gen6	2.0	$12 \pm 1$	N. A.
P411-gen6b	2.0	$18 \pm 1$	N. A.
P411-gen7	2.0	$34 \pm 1$	N. A.
P411-gen8	2.0	$130 \pm 20$	67.7 : 32.3
P411-gen9	2.0	$260 \pm 20$	71.9 : 28.1
P411-gen10	2.0	$510 \pm 30$	76.9 : 23.1
P411-gen11	2.0	$450 \pm 10$	83.0 : 17.0
P411-gen12	2.0	$630 \pm 20$	92.0 : 8.0
P411-gen13	2.0	$500 \pm 30$	96.9 : 3.1
P411-CHF	2.0	$440 \pm 30$	98.0 : 2.0

**Notes:** Reactions were performed using clarified *E. coli* lysate with 10 mM each substrate. Each reported TTN is the average over four reactions (performed from two independent cell cultures, each used for duplicate reactions). TTN, total turnover number, refers to TTN to **52i**; SD, standard deviation; N. A., not applicable, indicates that this value was not measured.

### C.5 Amino acid sequence of P411-CHF

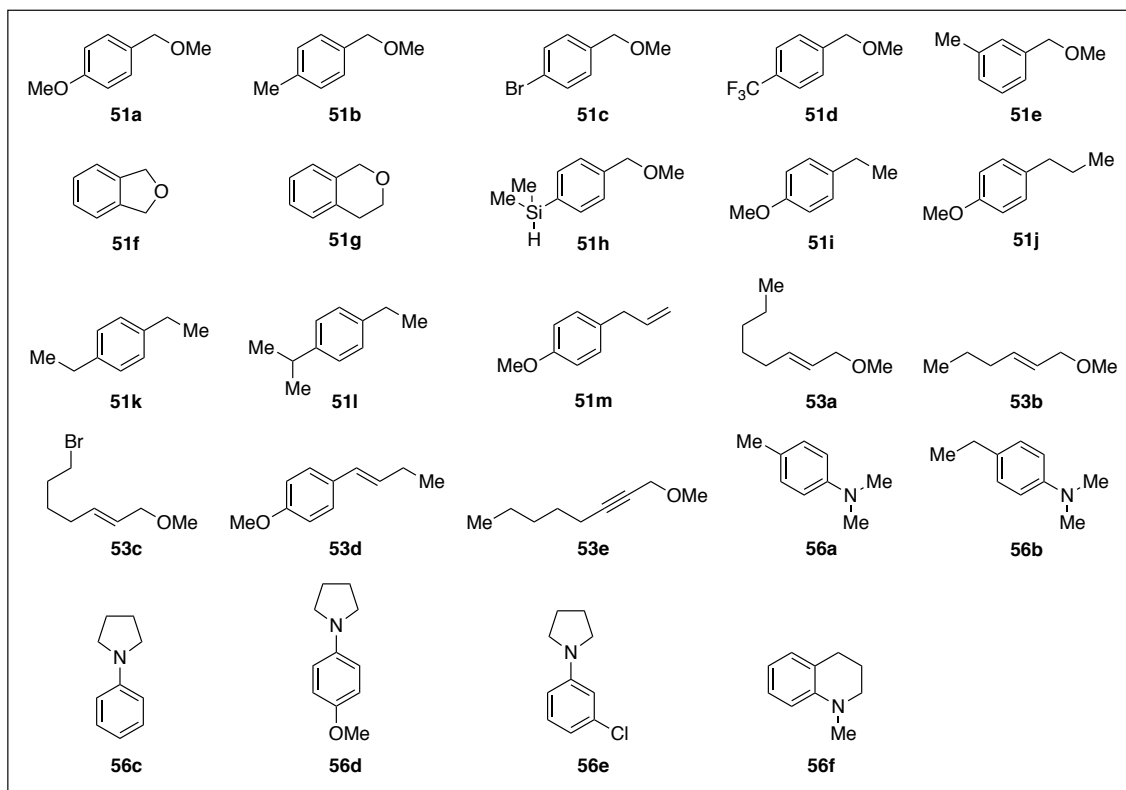
All heme proteins disclosed below were cloned into a pET22b(+) vector and contain a C-terminal 6xHis-tag.

Amino acid sequence of **P411-CHF**, an evolved C–H alkylation enzyme:

```
MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQRLIKEACDESRF
DKELSQPLKFLRDFLGDLATSWTHEKNWKKAHNILLPSFSQQAMKGYHAMMVDIAVQLVQKWERLNA
DEHIEVSEDMTRLTLDTIGLCGFNYRFNSFYRDQPHPFIIISLVRALDEV MNKLQRANPDDPAYDENKR
QFQEDIKVMNDLVDKIIADRKARGEQSDDLTLQMLNGKDPETGEPLDDGNIRYQIIITFLYAGVEGTSG
LLSFALYFLVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPTVPYFSLYAKEDT
VLGGEYPLEKGDEVMLIPQLHRDKTVWGDDVEEFRPERFENPSAIPQHAFKPFNGNGQRASIGQQFAL
HEATLVLGMMMLKHFD FEDHTNYELDIKELLTLKPKGFVVKAKSKKIPLGGIPSPSTEQS AKKVRKKAE
NAHNTPLLVLVYGSNMGTAE GTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPPDN
AKQFVDWLDQASADEVKGVRSVFGCGDKNWATTYQKVPAFIDETLAAKGAENIADRGEADASDDFEG
TYEEWREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADMPLAKMHGAFSTLEHHHHHH.
```

### C.5 Substrate synthesis and characterization

Commercially available alkane and diazo substrates were used as received: **51a**, **51d**, **51f**, **51g**, **51i–51k**, **51m**, **56a**, **56c–56f**, **58a**, **58e**, **58f** (custom synthesis, Arch Bioscience). Compound **1c** was also commercial (Combi-Blocks), though the commercial product was used only for synthesis. Ethyl diazoacetate (Sigma-Aldrich) was concentrated under reduced pressure and its concentration relative to residual dichloromethane was determined by  $^1\text{H}$  NMR. Diazo compounds **58h** (ref. 10) and **58i** (ref. 11) are known and were prepared according to literature procedures. Caution: although no safety issues were encountered, diazo compounds are reactive and should be used with caution.



**Figure C-10.** Alkane substrates. This terminology is used to emphasize that the reaction reported here acts on a  $sp^3$ -hybridized C–H bond in the alkane portion of the substrate.

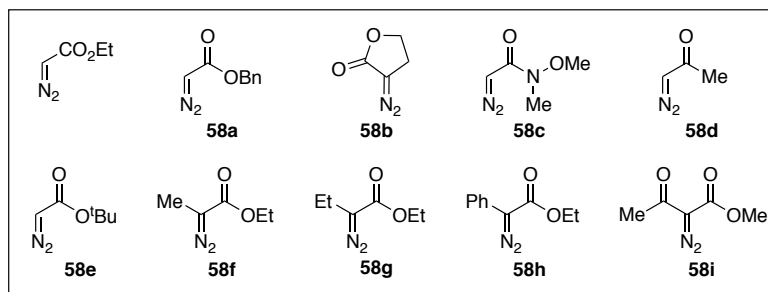


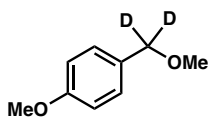
Figure C-11. Diazo compounds.

### General Procedure A: Methylation of alcohols

To a 250 mL round bottom flask was added NaH (60% dispersion in mineral oil, 15–30 mmol, 1.2–1.5 equiv.). The flask was evacuated and filled with argon (3 times). Anhydrous THF (45–80 mL) was added by syringe and the reaction mixture was cooled to 0 °C in an ice bath. Alcohol (10–20 mmol, 1.0 equiv.) in THF (5–10 mL) was added dropwise and the reaction mixture was allowed to warm to room temperature and stirred for 30 minutes. Following, iodomethane (20–40 mmol, 2.0 equiv.) in THF (10 mL) was added and the reaction was stirred at room temperature (8–15 hours). The reaction was quenched by the addition of brine (60 mL) or NH<sub>4</sub>Cl (sat. aq., 60 mL) and the phases were separated. The aqueous layer was extracted with diethyl ether (3 × 60 mL); the combined organics were washed with aq. sodium thiosulfate (10% w/v, 50 mL, when necessary), dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. Purification by silica column chromatography with hexanes / ethyl acetate or pentane / diethyl ether afforded the desired products in 37–99% yield.

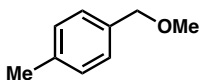
### 1-Methoxy-4-(methoxymethyl-*d*<sub>2</sub>)benzene (**51a-*d*<sub>2</sub>**)

Labeled substrate **51a-*d*<sub>2</sub>** was prepared from methyl 4-methoxybenzoate using a two-step sequence to 98% deuterium incorporation at the benzylic position. First, to a dry round bottom flask, under argon, was added LiAlD<sub>4</sub> (0.23 g, 5.5 mmol, 1.1 equiv.) and anhydrous Et<sub>2</sub>O (10 mL). A solution of methyl 4-methoxybenzoate (0.83 g, 5 mmol, 1.0 equiv.) in dry Et<sub>2</sub>O (5 mL) was added dropwise and the reaction was allowed to stir at room temperature for 12 hours. Following, the reaction mixture was cooled to 0 °C and diluted with Et<sub>2</sub>O. The reaction was quenched by the addition of 0.2 mL H<sub>2</sub>O, 0.2 mL NaOH (aq., 1M), and 0.6 mL H<sub>2</sub>O. The mixture was allowed to warm to room temperature and stirred for 15 minutes. MgSO<sub>4</sub> was added and the mixture was stirred for a further 15 minutes, filtered, and concentrated under reduced pressure. The crude product was purified by silica column chromatography with hexanes / ethyl acetate to give (4-methoxyphenyl)methanol-*d*<sub>2</sub> (0.43 g, 61% yield, 98% deuterium incorporation), with spectral data in agreement with literature report<sup>12</sup>. Methylation of this compound was performed following **General Procedure A** (note: reaction performed on 3.0 mmol scale) to afford **51a-*d*<sub>2</sub>** (0.43 g, 61% yield, 98% deuterium incorporation).



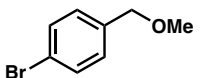
**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.27 (d, *J* = 8.4 Hz, 2H), 6.90 (d, *J* = 8.5 Hz, 2H), 3.81 (s, 3H), 3.36 (s, 3H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 159.3, 130.3, 129.5, 113.9, 73.7 (m, labeled), 57.8, 55.4. **HRMS** (EI) *m/z*: 154.0964 (*M*<sup>+</sup>); calc. for C<sub>9</sub>H<sub>10</sub>O<sub>2</sub>H<sub>2</sub>: 154.0963.

### 1-(Methoxymethyl)-4-methylbenzene (**51b**)



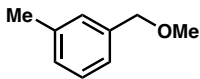
Prepared from *p*-tolylmethanol using **General Procedure A**. This compound is known in the literature<sup>13</sup>. **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.23 (d, *J* = 8.0 Hz, 2H), 7.17 (d, *J* = 7.9 Hz, 2H), 4.43 (s, 2H), 3.37 (s, 3H), 2.35 (s, 3H).

### 1-Bromo-4-(methoxymethyl)benzene (**51c**)



Prepared from (4-bromophenyl)methanol using **General Procedure A**. This compound is known<sup>13</sup>. **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.47 (d, *J* = 8.4 Hz, 2H), 7.21 (d, *J* = 8.6 Hz, 2H), 4.41 (s, 2H), 3.38 (s, 3H).

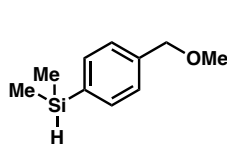
### 1-(Methoxymethyl)-3-methylbenzene (**51e**)



Prepared from *m*-tolylmethanol using **General Procedure A**. **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.28 – 7.21 (m, 1H), 7.19 – 7.08 (m, 3H), 4.43 (s, 2H), 3.40 (s, 3H), 2.36 (s, 3H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 138.2, 128.6, 128.5, 128.4, 124.9, 74.9, 58.3, 21.5. **HRMS** (FAB) *m/z*: 135.0810 [(*M* + *H*<sup>+</sup>) – H<sub>2</sub>]; calc. for C<sub>9</sub>H<sub>11</sub>O: 135.0810.

### (4-(Methoxymethyl)phenyl)dimethylsilane (**51h**)

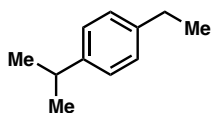
In a 250 mL round bottom flask, under argon, 1-bromo-4-(methoxymethyl)benzene (3.0 g, 15 mmol, 1.0 equiv.) in anhydrous THF (60 mL) was cooled to -78 °C. A solution of *n*-butyllithium (9 mL, 2.5 M in hexanes, 22.5 mmol, 1.5 equiv.) was added dropwise. The resulting mixture was stirred at -78 °C for 2 hours before the dropwise addition of chlorodimethylsilane (2.4 mL, 22.5 mmol, 1.5 equiv.). The reaction was allowed to warm to room temperature and stirred overnight. The reaction mixture was cooled to 0 °C and quenched with NH<sub>4</sub>Cl (sat. aq., 20 mL). The aqueous layer was extracted with diethyl ether (3 × 30 mL); the combined organics were washed with brine (30 mL), dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude reaction mixture was purified by silica column chromatography with hexanes / ethyl acetate to afford **51h** (2.14 g, 79% yield). A second round of purification by silica column chromatography with hexanes / ether was performed on a portion of the product.



**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.54 (d, *J* = 8.1 Hz, 2H), 7.35 (d, *J* = 8.1 Hz, 2H), 4.47 (s, 2H), 4.43 (hept, *J* = 3.7 Hz, 1H), 3.40 (s, 3H), 0.35 (d, *J* = 3.7 Hz, 6H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 139.3, 136.9, 134.3, 127.3, 74.7, 58.3, -3.6. **HRMS** (FAB) *m/z*: 179.0894 [(*M* + *H*<sup>+</sup>) - H<sub>2</sub>]; calc. for C<sub>10</sub>H<sub>15</sub>OSi: 179.0892.

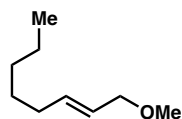
### 1-Ethyl-4-isopropylbenzene (**51l**)

The following procedure was modified from the literature<sup>14</sup>. To a 250 mL round bottom flask were added Pd/C (10% Pd on activated charcoal, 486 mg, 20% w/w), 4-isopropylacetophenone (2.43 g, 15 mmol), and methanol (60 mL). The solution was sparged with H<sub>2</sub> and stirred under 1 atm H<sub>2</sub> for 48 hours; monitoring the mixture by TLC showed that the reaction did not go to completion under these conditions. The crude reaction mixture was filtered through a pad of Celite, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. Purification by silica column chromatography with hexanes afforded product **51l** (218 mg, 1.47 mmol, 10% yield).



**<sup>1</sup>H NMR** (500 MHz, CDCl<sub>3</sub>) δ 7.19 – 7.13 (m, 4H), 2.90 (hept, *J* = 6.9 Hz, 1H), 2.65 (q, *J* = 7.6 Hz, 2H), 1.29 – 1.24 (m, 9H). **<sup>13</sup>C NMR** (126 MHz, CDCl<sub>3</sub>) δ 146.2, 141.7, 127.9, 126.5, 33.8, 28.5, 24.2, 15.7. **HRMS** (FAB) *m/z*: 149.1327 (*M* + *H*<sup>+</sup>); calc. for C<sub>11</sub>H<sub>17</sub>: 149.1330.

### (*E*)-1-Methoxyoct-2-ene (**53a**)

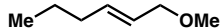


Prepared from (*E*)-oct-2-en-1-ol using **General Procedure A**. This compound is known in the literature<sup>15</sup>. **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 5.70 (dt, *J* = 15.6, 6.6, 1.2 Hz, 1H), 5.54 (dt, *J* = 15.3, 6.2, 1.4 Hz, 1H), 3.86 (dq, *J* = 6.2, 1.0 Hz, 2H), 3.31 (s, 3H), 2.08 – 1.99 (m, 2H), 1.43 – 1.34 (m, 2H), 1.34 – 1.21 (m, 4H), 0.88 (t, *J* = 7.0 Hz, 3H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 135.2, 126.1, 73.5, 57.8, 32.4, 31.5, 28.9, 22.7, 14.2.

### (*E*)-1-Methoxyhex-2-ene (**53b**)

Prepared from (*E*)-hex-2-en-1-ol using a modified version of **General Procedure A**. To a 100 mL dry round bottom flask, cooled under argon, were added (*E*)-hex-2-en-1-ol (2.0 g, 20 mmol, 1.0 equiv.), DMF (35 mL), and iodomethane (5.7 g, 40 mmol, 2.0 equiv.). The resulting solution was cooled to 0 °C and NaH (60% dispersion in mineral oil, 960 mg, 24 mmol, 1.2 equiv.) was added portion-wise. The mixture was stirred at 0 °C for 30 minutes, then allowed to warm to room temperature and stirred for an additional 3 hours. The reaction mixture was cooled to 0 °C, quenched with the addition of NH<sub>4</sub>Cl (sat. aq., 30 mL), and diluted with diethyl ether (50 mL). Phases were separated and the aqueous layer was extracted with diethyl ether (3 × 50 mL). The combined organics were washed with H<sub>2</sub>O (2 × 25 mL) and brine (25 mL), dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure (≥ 200 mbar). Purification by silica column chromatography with pentane

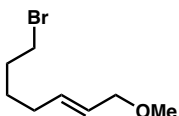
/ diethyl ether afforded compound **53b** (746 mg, 6.5 mmol, 33% yield).



This compound is known in the literature<sup>16</sup>. <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 5.70 (dt, *J* = 15.4, 6.6, 1.2 Hz, 1H), 5.55 (dt, *J* = 15.4, 6.3, 1.4 Hz, 1H), 3.87 (dq, *J* = 6.3, 1.1 Hz, 2H), 3.32 (s, 3H), 2.06 – 2.00 (m, 2H), 1.42 (*app.* sext, *J* = 7.4 Hz, 2H), 0.91 (t, *J* = 7.3 Hz, 3H). <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>) δ 134.9, 126.3, 73.4, 57.8, 34.5, 22.4, 13.8.

### (*E*)-7-Bromo-1-methoxyhept-2-ene (**53c**)

To a 100 mL flamed dried flask was added Grubbs' catalyst 2<sup>nd</sup> generation (85 mg, 1 mol%). The flask was then evacuated and backfilled with argon for three times. Under argon, a dry CH<sub>2</sub>Cl<sub>2</sub> solution containing 6-bromo-1-hexene (1.63 g, 10 mmol, 1.0 equiv.) and crotonaldehyde (3.50 g, 50 mmol, 5.0 equiv.) was added to the flask. The mixture was stirred under reflux for 20 hours and then cooled to room temperature and filtered through a silica plug. The solvent was removed under reduced pressure and the crude product was purified by flash chromatography (hexanes / ethyl acetate) to give (*E*)-7-bromohept-2-enal (1.6 g, 84% yield). This product was then dissolved in 10 mL dry THF and then added slowly to a suspension of NaBH<sub>4</sub> (375 mg, 10 mmol, 1.0 equiv.) in dry THF (10 mL) at 0 °C. To this reaction mixture, iodine (1.27 g, 5 mmol, 0.5 equiv.) in 10 mL of THF was slowly added at 0 °C. Reaction was stirred until the aldehyde was fully reduced as indicated by TLC. The reaction was quenched with NH<sub>4</sub>Cl (sat. aq.), the phases were separated, and the aqueous phase was extracted with ethyl acetate (3 × 20 mL). The combined organic layers were washed with brine and dried over Na<sub>2</sub>SO<sub>4</sub>. The solvent was removed under reduced pressure and the crude alcohol product was used directly without purification. **General Procedure A** was used for the methylation step and the final product **53c** was obtained with 50% overall yield (1.03g, 5 mmol). This compound was prepared by Dr. Xiongyi Huang.



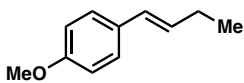
This compound is known in the literature<sup>17</sup>. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 5.68 (dt, *J* = 15.3, 6.4, 1.1 Hz, 1H), 5.57 (dt, *J* = 15.4, 6.0, 1.2 Hz, 1H), 3.86 (dq, *J* = 5.9, 1.0 Hz, 2H), 3.41 (t, *J* = 6.8 Hz, 2H), 3.32 (s, 3H), 2.14 – 2.05 (m, 2H), 1.92 – 1.82 (m, 2H), 1.57 – 1.49 (m, 2H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 133.9, 127.0, 73.3, 57.9, 33.8, 32.3, 31.5, 27.7. HRMS (EI) *m/z*: 205.0216 (*M* – H<sup>+</sup>); calc. for C<sub>8</sub>H<sub>14</sub><sup>79</sup>BrO: 205.0228.

### (*E*)-1-(But-1-en-1-yl)-4-methoxybenzene (**53d**)

This compound was accessed in a two-step sequence. First, to propyltriphenylphosphonium bromide (7.6 g, 19.7 mmol, 1.0 equiv.) suspended in anhydrous THF (70 mL) and cooled to 0 °C was added *n*-butyllithium (2.5 M in hexanes, 7.9 mL, 19.7 mmol, 1.0 equiv.) dropwise over 10 min to form a bright orange solution. After stirring for 1 hour, 4-methoxybenzaldehyde (2.7 g, 19.7 mmol, 1.0 equiv.) was added dropwise over 5 min. The reaction mixture was allowed to slowly warm to room temperature and stirred at room temperature overnight. The reaction mixture was diluted

with pentane (50 mL) and the resulting solution was washed with HCl (aq., 0.1 M, 50 mL), H<sub>2</sub>O (50 mL), NaHCO<sub>3</sub> (sat. aq., 50 mL), and brine (50 mL). The organics were dried over sodium sulfate and concentrated under reduced pressure. The crude product was purified by silica column chromatography with pentane / diethyl ether to afford (*E:Z*)-**53d** (2:1 *E:Z*, 2.50 g, 15.4 mmol, 78% yield).

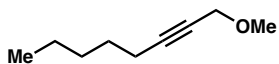
Next, (*E:Z*)-**53d** was isomerized following a literature method<sup>18</sup>. To a dry 25 mL round bottom flask, under argon, were added (*E:Z*)-**53d** (300 mg, 1.85 mmol), (MeCN)<sub>2</sub>PdCl<sub>2</sub> (235 mg, 50 mol%), and 4 mL anhydrous dichloromethane. The resulting mixture was stirred at room temperature for 24 hours. The crude reaction mixture was filtered through Celite and concentrated under reduced pressure. Purification by silica column chromatography using hexanes / diethyl ether delivered **53d** (> 20:1 *E:Z*, 279 mg, 1.72 mmol, 93% yield).



This compound is known in the literature<sup>18</sup>. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.28 (d, *J* = 8.7 Hz, 2H), 6.85 (d, *J* = 8.8 Hz, 2H), 6.33 (dt, *J* = 15.7, 1.6 Hz, 1H), 6.13 (dt, *J* = 15.8, 6.5 Hz, 1H), 3.80 (s, 3H), 2.26 – 2.16 (m, 2H), 1.08 (t, *J* = 7.5 Hz, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 158.7, 130.9, 130.7, 128.2, 127.1, 114.0, 55.4, 26.2, 14.0.

#### 1-Methoxyoct-2-yne (53e)

To a solution of 3-methoxyprop-1-yne (845 μL, 10 mmol, 1.0 equiv.) in anhydrous THF (50 mL) at -20 °C, was added *n*-butyllithium (2 M in THF, 6 mL, 12 mmol, 1.2 equiv.) and HMPA (869 μL, 5 mmol, 0.5 equiv.) dropwise over 5 min. The resulting mixture was stirred at -20 °C for 3 hours before the addition of 1-iodopentane (1.96 mL, 15 mmol, 1.5 equiv.). The reaction was allowed to slowly warm to room temperature in 2 hours and stirred for additional 18 hours. The reaction was then quenched by NH<sub>4</sub>Cl (sat. aq., 20 mL) and H<sub>2</sub>O (30 mL), and extracted by diethyl ether (40 mL × 3). The combined organic layer was washed by H<sub>2</sub>O (50 mL) and brine (50 mL), and then dried over sodium sulfate and concentrated under reduced pressure. The crude product was purified by silica column chromatography with pentane / diethyl ether to afford **53e** (1.04 g, 7.4 mmol, 74% yield).



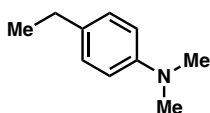
This compound is known in the literature<sup>19</sup>. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.07 (t, *J* = 2.2 Hz, 2H), 3.37 (s, 3H), 2.22 (tt, *J* = 7.2, 2.2 Hz, 2H), 1.56 – 1.47 (m, 2H), 1.41 – 1.26 (m, 4H), 0.89 (t, *J* = 7.1 Hz, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 87.4, 75.8, 60.4, 57.5, 31.2, 28.5, 22.3, 18.9, 14.1. HRMS (EI) *m/z*: 139.1128 [(M – H)<sup>+</sup>]; calc. for C<sub>9</sub>H<sub>15</sub>O: 139.1123.

#### 4-Ethyl-*N,N*-dimethylaniline (56b)

4-Ethylaniline (0.605 g, 5 mmol, 1.0 equiv.) and formaldehyde (1.8 mL, 50 mmol, 10.0 equiv.) were mixed in acetic acid (30 mL). The solution was stirred for 30 min at 0 °C before portionwise addition of NaBH<sub>3</sub>CN (1.57 g, 25 mmol, 5.0 equiv.). After the reaction was stirred overnight, NaOH (aq., 2M) was used to neutralize the reaction at 0 °C until pH 8-10. The crude product was extracted with diethyl ether (30 mL × 3). The

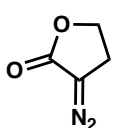


combined organic layer was washed with H<sub>2</sub>O (50 mL) and brine (50 mL), and then dried over sodium sulfate and concentrated under reduced pressure. The crude product was purified by silica column chromatography with hexanes / ethyl acetate to afford **56b** (635 mg, 4.25 mmol, 85% yield).



This compound is known in the literature<sup>20</sup>. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.09 (d, *J* = 8.8 Hz, 2H), 6.72 (d, *J* = 8.7 Hz, 2H), 2.92 (s, 6H), 2.57 (q, *J* = 7.6 Hz, 2H), 1.21 (t, *J* = 7.6 Hz, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 149.1, 132.8, 128.5, 113.3, 41.2, 27.9, 16.1.

### 3-Diazodihydrofuran-2(3*H*)-one (**58b**)



The preparation of the title compound **58b** followed a modified procedure reported by Sattely *et al*<sup>21</sup>. Sodium azide (4.83 g, 74.3 mmol, 4 equiv.), sodium hydroxide (80 mL of 2 M in water, 160 mmol), tetrabutylammonium bromide (60.0 mg, 0.190 mmol, 0.01 equiv.), and hexane (80 mL) were combined in a 500-mL flask with magnetic stir bar open to the air and cooled to 0 °C. With vigorous stirring, triflic anhydride (6.20 mL, 37.1 mmol, 2 equiv.) was added dropwise. After 15 min, a solution of 2-acetyl-butylolactone (2.00 mL, 18.6 mmol) in acetonitrile (70 mL) was poured into the vessel through a funnel, followed by an additional 10 mL of acetonitrile to complete the transfer. The initially colorless reaction mixture immediately turned yellow. After stirring for 20 min at 0 °C, the mixture was diluted with ice water (50 mL) and chilled EtOAc (50 mL) and transferred to a separatory funnel. After phase separation and removal of the organic fraction, the aqueous layer was washed with chilled EtOAc (50 mL × 5). The combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated under reduced pressure. The resulting crude product was purified by silica column chromatography with hexanes / ethyl acetate as eluents. The yellow-colored fractions were concentrated to afford the product as a bright yellow crystalline solid (1.2–1.6 g, 60–75% yield). Spectral data are consistent with Sattely *et al*<sup>21</sup>. This compound was prepared by Kai Chen.

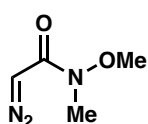
### 2-Diazo-*N*-methoxy-*N*-methylacetamide (**58c**)

4-Methylbenzenesulfonohydrazide (9.31 g, 50 mmol, 1.0 equiv.) was dissolved in aqueous hydrochloric acid (2 M, 30 mL) and warmed to 50 °C (solution 1). 2-Oxoacetic acid (7.40 g of 50% in water, 50 mmol, 1.0 equiv.) was dissolved in water (100 mL) and heated to 50 °C (solution 2). Pre-warmed solution 1 was slowly transferred to solution 2. The reaction mixture was then stirred at 60 °C for 4 h until all the hydrozone product crashed out. The mixture was cooled to 0 °C and kept for 2 h. The product 2-(2-tosylhydrazineylidene)acetic acid (9.88 g, 82% yield) was collected by filtration, washed with hexane: ethyl acetate (10:1) and dried under vacuum.

2-(2-Tosylhydrazineylidene)acetic acid (4.84 g, 20 mmol, 1.0 equiv.) was dissolved in dry dichloromethane (30 mL). Thionyl chloride (16 mL) and *N,N*-dimethyl formaldehyde (3 drops, *cat.*) were added to the solution. The reaction mixture was stirred

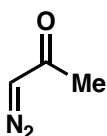
at room temperature for 1 h and then heated to reflux ( $\sim 50\text{ }^{\circ}\text{C}$ ) for 5 h until the starting material was completely dissolved and the reaction turned clear and light yellow. After the reaction was cooled to room temperature, organic solvent and the excess thionyl chloride was removed under reduced pressure. The resulting mixture (solid) was then dissolved in dry dichloromethane (20 mL) and used for the next step.

*N,O*-Dimethylhydroxylamine hydrochloride (3.91 g, 40 mmol, 2.0 equiv.) and triethylamine (11.2 mL, 80 mmol, 4.0 equiv.) were mixed in dry dichloromethane (80 mL) and stirred for 30 min. The solution of acyl chloride was added dropwise over 20 min to the reaction mixture at  $0\text{ }^{\circ}\text{C}$ . The reaction was then stirred at room temperature for 5 h before water (80 mL) was added to quench the reaction. The liquid phases were transferred to a separatory funnel, and the aqueous phase was extracted with dichloromethane ( $50\text{ mL} \times 4$ ). The combined organic phase was washed with water (40 mL) and brine (40 mL), dried over  $\text{Na}_2\text{SO}_4$ , filtered, and concentrated under reduced pressure. The resulting crude product was purified by silica column chromatography with hexanes / ethyl acetate to afford **9c** as a yellow liquid (0.82 g, 32% yield). This compound was prepared by Kai Chen.



This compound is known in the literature<sup>22</sup>.  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  5.33 (s, 1H), 3.66 (s, 3H), 3.19 (s, 3H).

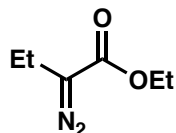
### 1-Diazopropan-2-one (**58d**)



The preparation of the title compound **9d** followed a modified procedure reported by Zhang *et al.*<sup>23</sup>. To a solution of acetylacetone (3.4 mL, 33.0 mmol, 1.10 equiv.) and triethylamine (5.04 mL, 36.4 mmol, 1.21 equiv.) in dry acetonitrile (25 mL), a solution of *p*-acetamidobenzenesulfonyl azide (7.20 g, 30.0 mmol, 1.0 equiv.) in dry acetonitrile (25 mL) was added dropwise. The reaction mixture was stirred at room temperature for 4 h. Then, the solvent was removed under reduced pressure and the resulting mixture was then purified by silica column chromatography with hexanes / ethyl acetate to give 3-diazopentane-2,4-dione (3.65 g, 96% yield) as a pale yellow liquid.

3-Diazopentane-2,4-dione (1.89 g, 15 mmol, 1.0 equiv.) was dissolved in diethyl ether (25 mL). An aqueous solution (25 mL) of NaOH (3.00 g, 75 mmol, 5.0 equiv.) was added dropwise over 10 min to the ether layer with vigorous stirring at  $0\text{ }^{\circ}\text{C}$ . The reaction mixture turned dark brown within 20 min and was then stirred at room temperature for 4 h. The liquid phases were transferred to a separatory funnel, and the aqueous phase was extracted with diethyl ether ( $30\text{ mL} \times 5$ ). The combined organic phase was dried over  $\text{Na}_2\text{SO}_4$ , filtered, and concentrated under reduced pressure ( $T = 24\text{ }^{\circ}\text{C}$ ,  $P \geq 20\text{ kPa}$ ) to give product **58d** as a volatile yellow liquid (0.892 g, 71% yield). Spectral data is consistent with Zhang *et al.*<sup>23</sup>. This compound was prepared by Kai Chen and Dr. Xiongyi Huang.

### Ethyl 2-diazobutanoate (**58g**)



The preparation of the title compound **58g** followed a modified procedure reported by Huang *et al*<sup>24</sup>. To a solution of ethyl 2-ethylacetoacetate (3.16 g, 20.0 mmol, 1.0 equiv.) and *p*-acetamidobenzenesulfonyl azide (7.21 g, 30.0 mmol, 1.5 equiv.) in dry acetonitrile (50 mL) was added 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU, 4.5 mL, 30.0 mmol, 1.5 equiv.) dropwise at 0 °C. The reaction mixture was stirred at 0 °C for 1 h and at room temperature for 2 h. Water (50 mL) was added to quench the reaction. Acetonitrile was removed under reduced pressure (T = 24 °C, P ≥ 20 kPa). The mixture was extracted with diethyl ether (25 mL × 4). The combined ether layer extract was washed with water (30 mL) and brine (30 mL), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated under reduced pressure (T = 24 °C, P ≥ 30 kPa). The crude product was purified by silica column chromatography with hexanes / ethyl acetate to give product **9g** as a volatile yellow liquid (2.40 g, 84% yield). Spectral data is consistent with Huang *et al*<sup>24</sup>. This compound was prepared by Kai Chen.

## C.6 Synthesis and characterization of reference compounds

Racemic reference compounds corresponding to enzymatic products and side-products were prepared according to the following procedures. Reference compounds are characterized below.

### General Procedure B: Aldol reaction and methylation synthetic sequence

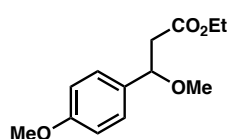
In a dry 100 or 250 mL round bottom flask, under argon, a solution of diisopropylamine (6–24 mmol, 1.1–1.2 equiv.) in THF (15–30 mL) was cooled to 0 °C (**General Procedure B-1**) or -78 °C (**General Procedure B-2**). *n*-Butyllithium (6–25 mmol, 1.1–1.2 equiv., 1.6 or 2.5 M in hexanes) was added dropwise and the resulting mixture was stirred at the appropriate temperature for 15–30 min. The mixture was cooled to -78 °C and kept at this temperature for the remainder of the reaction. Ethyl acetate (14–28 mmol, 1.4 equiv., **General Procedure B-1** or 6–10 mmol, 1.0 equiv., **General Procedure B-2**) was added dropwise and the mixture was stirred for an additional 30–45 min. Then, aldehyde (10–20 mmol, 1.0 equiv., **General Procedure B-1** or 9–11 mmol, 1.1–1.5 equiv., **General Procedure B-2**) as a solution in THF (15–30 mL, **General Procedure B-1**) or neat (**General Procedure B-2**) was added slowly and the solution was stirred for a further 0.5–3 hours. The reaction mixture was quenched at -78 °C by the addition of NH<sub>4</sub>Cl (sat. aq., 10–30 mL) and allowed to thaw to room temperature. For **General Procedure B-1** only, HCl (1 M aq., 1.5–3.0 mL) was also added. Phases were separated and the aqueous phase was extracted with ethyl acetate or diethyl ether (3 × 20–30 mL). The combined organics were washed with NH<sub>4</sub>Cl (sat. aq., 2 × 10–15 mL), brine (10 mL), dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. Purification by silica column chromatography with hexanes / ethyl acetate afforded the desired aldol adducts in 56–95% yield.

In the appropriate reaction vessel, aldol adduct (3–4 mmol, 1.0 equiv.), Ag<sub>2</sub>O (9–10 mmol, 2.5–3.0 equiv.), and solvent (10–15 mL) were combined, followed by iodomethane (40–60 mmol, 10–15 equiv., **General Procedure B-1** or 9 mmol, 3.0 equiv., **General Procedure B-2**). The reaction was then stirred at the specified temperature for 24–48 hours, with additional equivalents of iodomethane (10–20 mmol, 2.5–5 equiv., **General Procedure B-1**) added as necessary. For **General Procedure B-1**, the reaction was performed in a vial equipped with a pressure release cap, toluene was employed as the solvent, and the reaction mixture was stirred at 70 °C. For **General Procedure B-2**, diethyl ether was employed as solvent and the reaction mixture was stirred at room temperature; the reaction vessel was covered in aluminum foil to protect its contents from light. The crude mixture was filtered through a pad of Celite and concentrated under reduced pressure. Purification was performed by silica column chromatography with hexanes / ethyl acetate; if necessary, a second purification by reverse phase chromatography was performed (Biotage Isolera equipped with Biotage SNAP Ultra C18 column, water / acetonitrile eluent system). The desired products were obtained in 25–57% yield.

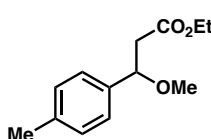
### **General Procedure C: Horner-Wadsworth-Emmons reaction and Pd/C alkene reduction synthetic sequence**

In a dry round bottom flask, under argon, NaH (60% dispersion in mineral oil, 7.4–12 mmol, 1.1–2.0 equiv.) in anhydrous THF (8–23 mL) was cooled to 0 °C. Triethyl phosphonoacetate (7.4–18 mmol, 1.1–3.0 equiv.) was added dropwise and the mixture was allowed to warm to room temperature and stirred for 1 hour. Ketone (5–6.7 mmol, 1.0 equiv.) in THF (2–4 mL) was added and the reaction was stirred at room temperature for 12–18 hours (for the preparation of **3j** and **3l**) or heated to reflux (for the preparation of **3i**, **3k**, **8a'**, and **8b'**). The reaction was quenched with NH<sub>4</sub>Cl (sat. aq., 20 mL). Phases were separated and the aqueous layer was extracted with ethyl acetate (3 × 30 mL). The combined organics were washed with brine (10–20 mL), dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. When necessary, the crude product was purified by silica column chromatography with hexanes / ethyl acetate to afford the desired alkene compounds in 23% to quantitative yield.

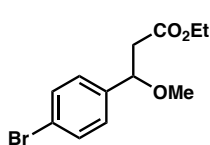
To a round bottom flask were added Pd/C (10% Pd on activated charcoal, 24–30% w/w of alkene), methanol (5–6 mL), and alkene (1.2–2.3 mmol). H<sub>2</sub> was bubbled through the solution for ~30 minutes. The reaction was stirred at room temperature under 1 atm H<sub>2</sub> until complete reduction of the alkene was observed by TLC (typically 3–8 hours). The crude product was filtered through a pad of Celite and concentrated under reduced pressure. Purification by silica column chromatography with hexanes / ethyl acetate afforded the desired products in quantitative yield.

**Ethyl 3-methoxy-3-(4-methoxyphenyl)propanoate (52a)**

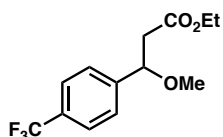
This compound was prepared from 4-methoxybenzaldehyde using **General Procedure B-1**.  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.25 (d,  $J = 8.5$  Hz, 2H), 6.89 (d,  $J = 8.8$  Hz, 2H), 4.58 (dd,  $J = 9.0, 4.9$  Hz, 1H), 4.14 (qd,  $J = 7.1, 1.2$  Hz, 2H), 3.81 (s, 3H), 3.19 (s, 3H), 2.80 (dd,  $J = 15.2, 9.0$  Hz, 1H), 2.55 (dd,  $J = 15.2, 4.9$  Hz, 1H), 1.23 (t,  $J = 7.1$  Hz, 3H). Spectral data are in agreement with that for the enzymatic product (see Section C.8).

**Ethyl 3-methoxy-3-(*p*-tolyl)propanoate (52b)**

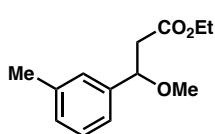
This compound was prepared from 4-methylbenzaldehyde using **General Procedure B-1**.  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.22 (d,  $J = 8.0$  Hz, 2H), 7.16 (d,  $J = 7.9$  Hz, 2H), 4.60 (dd,  $J = 9.2, 4.7$  Hz, 1H), 4.14 (qd,  $J = 7.1, 1.2$  Hz, 2H), 3.21 (s, 3H), 2.79 (dd,  $J = 15.3, 9.2$  Hz, 1H), 2.55 (dd,  $J = 15.3, 4.7$  Hz, 1H), 2.35 (s, 3H), 1.24 (t,  $J = 7.2$  Hz, 3H).  $^{13}\text{C NMR}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  171.2, 137.9, 137.6, 129.4, 126.7, 80.0, 60.7, 56.9, 43.7, 21.3, 14.3. **HRMS** (FAB)  $m/z$ : 221.1169  $[(M + H^+) - H_2]$ ; calc. for  $\text{C}_{13}\text{H}_{17}\text{O}_3$ : 221.1178.

**Ethyl 3-(4-bromophenyl)-3-methoxypropanoate (52c)**

This compound was prepared from 4-bromobenzaldehyde using **General Procedure B-1**.  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.48 (d,  $J = 8.4$  Hz, 2H), 7.22 (d,  $J = 8.3$  Hz, 2H), 4.59 (dd,  $J = 8.9, 5.0$  Hz, 1H), 4.14 (qd,  $J = 7.1, 0.7$  Hz, 2H), 3.21 (s, 3H), 2.77 (dd,  $J = 15.4, 8.9$  Hz, 1H), 2.53 (dd,  $J = 15.4, 5.0$  Hz, 1H), 1.23 (t,  $J = 7.1$  Hz, 3H).  $^{13}\text{C NMR}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  170.8, 139.8, 131.9, 128.5, 122.0, 79.6, 60.8, 57.1, 43.5, 14.3. **HRMS** (FAB)  $m/z$ : 287.0282 ( $M + H^+$ ); calc. for  $\text{C}_{12}\text{H}_{16}^{79}\text{BrO}_3$ : 287.0283

**Ethyl 3-methoxy-3-(4-(trifluoromethyl)phenyl)propanoate (52d)**

This compound was prepared from 4-(trifluoromethyl)benzaldehyde using **General Procedure B-1**.  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.62 (d,  $J = 8.0$  Hz, 2H), 7.46 (d,  $J = 8.2$  Hz, 2H), 4.70 (dd,  $J = 8.8, 4.9$  Hz, 1H), 4.15 (qd,  $J = 7.2, 0.7$  Hz, 2H), 3.24 (s, 3H), 2.79 (dd,  $J = 15.5, 8.9$  Hz, 1H), 2.56 (dd,  $J = 15.4, 4.9$  Hz, 1H), 1.24 (t,  $J = 7.1$  Hz, 3H).  $^{13}\text{C NMR}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  170.7, 145.0, 130.4 (q,  $J = 32.4$  Hz), 127.1, 125.7 (q,  $J = 3.8$  Hz), 124.2 (q,  $J = 272.1$  Hz), 79.7, 60.9, 57.3, 43.5, 14.3. **HRMS** (FAB)  $m/z$ : 277.1041 ( $M + H^+$ ); calc. for  $\text{C}_{13}\text{H}_{16}\text{F}_3\text{O}_3$ : 277.1052

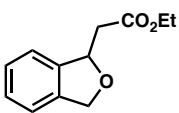
**Ethyl 3-methoxy-3-(*m*-tolyl)propanoate (52e)**

This compound was prepared from 3-methylbenzaldehyde using **General Procedure B-1**.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.27 – 7.21 (m, 1H), 7.16 – 7.08 (m, 3H), 4.60 (dd,  $J = 9.2, 4.6$  Hz, 1H), 4.15 (qd,  $J = 7.1, 1.3$  Hz, 2H), 3.22 (s, 3H), 2.79 (dd,  $J = 15.3, 9.3$  Hz, 1H), 2.56 (dd,  $J = 15.3, 4.6$  Hz, 1H), 2.36 (s, 3H), 1.24 (t,  $J = 7.1$  Hz, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  171.2, 140.7, 138.3, 128.9, 128.6, 127.4, 123.8, 80.2, 60.7, 57.0, 43.7, 21.6, 14.3. HRMS (FAB)  $m/z$ : 223.1338 ( $\text{M} + \text{H}^+$ ); calc. for  $\text{C}_{13}\text{H}_{19}\text{O}_3$ : 223.1334.

**Ethyl 2-(1,3-dihydroisobenzofuran-1-yl)acetate (52f)**

This compound was prepared by the method of U. S. Dakarapu *et al.*<sup>25</sup>. To a flame-dried Schlenk flask under argon was added  $[\text{Ir}(\text{coe})_2\text{Cl}]_2$  (5 mg, 0.0056 mmol, 0.11 mol%), phthalide (671 mg, 5 mmol, 1.0 equiv.), anhydrous dichloromethane (1.6 mL), and  $\text{H}_2\text{SiEt}_2$  (1.3 mL, 10 mmol, 2 equiv.). The reaction mixture was stirred for 14 hours at room temperature. The reaction mixture was concentrated under reduced pressure to afford the crude silyl acetal, which was used without purification.

In a dry round bottom flask, the crude silyl acetal (5 mmol, 1.0 equiv.) was combined with THF (5 mL) and the resulting mixture cooled to 0 °C. To the mixture were added triethyl phosphonoacetate (1.23 g, 5.5 mmol, 1.1 equiv.) and  $\text{KOSiMe}_3$  (713 mg, 5 mmol, 1.0 equiv.) in THF (7.5 mL). The reaction was allowed to warm to room temperature and stirred for 1.5 hours. The reaction was quenched with the addition of  $\text{NH}_4\text{Cl}$  (sat. aq., 12 mL) and the aqueous phase was extracted with diethyl ether ( $3 \times 15$  mL). The combined organics were washed with brine (15 mL), dried over  $\text{Na}_2\text{SO}_4$ , and concentrated under reduced pressure. Purification by silica column chromatography with hexanes / ethyl acetate afforded desired product **3f** with impurities (667 mg, 3.2 mmol, 65% yield). A portion of the product was taken for a second purification by reverse phase chromatography (Biotage Isolera equipped with Biotage SNAP Ultra C18 column, water / acetonitrile eluent system).

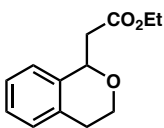


Spectral data are in agreement with literature report<sup>25</sup>.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.33 – 7.26 (m, 2H), 7.25 – 7.16 (m, 2H), 5.71 – 5.63 (m, 1H), 5.19 – 5.13 (m, 1H), 5.11 – 5.04 (m, 1H), 4.20 (q,  $J = 7.1$  Hz, 2H), 2.80 (dd,  $J = 15.6, 4.9$  Hz, 1H), 2.73 (dd,  $J = 15.6, 7.9$  Hz, 1H), 1.27 (t,  $J = 7.1$  Hz, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  171.0, 140.8, 139.3, 128.0, 127.6, 121.3, 121.2, 80.5, 72.9, 60.8, 41.8, 14.3.

**Ethyl 2-(isochroman-1-yl)acetate (52g)**

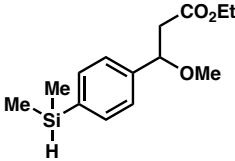
This compound was prepared by the method of R. E. TenBrink *et al.*<sup>26</sup>. To a 100 mL dry round bottom flask, under argon, were added 2-phenylethanol (1.47 g, 12 mmol, 1.0 equiv.), ethyl 3,3-diethoxypropionate (90% technical grade, 2.51 g, 13.2 mmol, 1.1 equiv.),

and anhydrous dichloromethane (5 mL). The resulting mixture was cooled to 0 °C and TiCl<sub>4</sub> (1 M in dichloromethane, 26.4 mL, 26.4 mmol, 2.2 equiv.) was added slowly. The reaction was stirred for 2 hours at 0 °C and a second portion of ethyl 3,3-diethoxypropionate (90% technical grade, 0.12 g, 0.6 mmol, 0.05 equiv.) was added. The reaction was stirred for an additional 2 hours at 0 °C. The mixture was poured into ice cold HCl (aq., 1 M, 20 mL) and the aqueous phase was extracted with dichloromethane (2 x 20 mL). The combined organics were washed with brine (30 mL), dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. Purification by silica column chromatography with hexanes / ethyl acetate afforded desired product **3g** with minor impurities (2.59 g, ~11.8 mmol, ~98% yield). A portion of the product was taken for a second purification by reverse phase chromatography (Biotage Isolera equipped with Biotage SNAP Ultra C18 column, water / acetonitrile eluent system).

 Spectral data are in agreement with literature report<sup>26</sup>. **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.22 – 7.15 (m, 2H), 7.15 – 7.09 (m, 1H), 7.08 – 7.02 (m, 1H), 5.25 (dd, *J* = 9.6, 3.5 Hz, 1H), 4.22 (q, *J* = 7.1 Hz, 2H), 4.13 (ddd, *J* = 11.4, 5.2, 4.2 Hz, 1H), 3.82 (ddd, *J* = 11.4, 9.0, 3.9 Hz, 1H), 3.04 – 2.93 (m, 1H), 2.88 (dd, *J* = 15.2, 3.6 Hz, 1H), 2.80 – 2.68 (m, 2H), 1.28 (t, *J* = 7.2 Hz, 3H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 171.4, 136.9, 134.1, 129.2, 126.8, 126.4, 124.6, 73.1, 63.2, 60.8, 41.9, 28.9, 14.3.

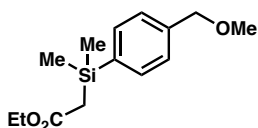
### Ethyl 3-(4-(dimethylsilyl)phenyl)-3-methoxypropanoate (**52h**)

This compound was prepared from ethyl 3-(4-bromophenyl)-3-methoxypropanoate (**3c**). The following procedure was modified from the literature<sup>27</sup>. To a 25 mL round bottom flask was added Mg turnings\* (48 mg, 2.0 mmol, 2.0 equiv.), flame dried, and cooled under positive argon pressure. (\*Mg turnings were prepared by washing with 0.1 M HCl, sonication, then washing with H<sub>2</sub>O and acetone.) THF (3 mL), LiCl (64 mg, 1.5 mmol, 1.5 equiv.), and Me<sub>2</sub>SiHCl (170 mg, 1.8 mmol, 1.8 equiv.) were added and the resulting mixture was stirred for 30 minutes at room temperature under positive argon pressure. Aryl bromide **52c** (287 mg, 1.0 mmol, 1.0 equiv.) was added dropwise via syringe and the reaction was stirred for an additional 2 hours. The crude reaction mixture was filtered through a pad of Celite and concentrated under reduced pressure. Purification by silica column chromatography with hexanes / ethyl acetate afforded desired product **52h** (145 mg, 0.54 mmol, 54% yield).

 **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.54 (d, *J* = 8.0 Hz, 2H), 7.33 (d, *J* = 7.9 Hz, 2H), 4.63 (dd, *J* = 9.3, 4.5 Hz, 1H), 4.42 (hept, *J* = 3.8 Hz, 1H), 4.15 (qd, *J* = 7.1, 1.6 Hz, 2H), 3.23 (s, 3H), 2.79 (dd, *J* = 15.3, 9.3 Hz, 1H), 2.56 (dd, *J* = 15.3, 4.5 Hz, 1H), 1.24 (t, *J* = 7.2 Hz, 3H), 0.35 (d, *J* = 3.8 Hz, 6H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 171.1, 141.8, 137.4, 134.4, 126.2, 80.2, 60.7, 57.1, 43.7, 14.3, -3.6. **HRMS** (FAB) *m/z*: 265.1253 [(M + H<sup>+</sup>) - H<sub>2</sub>]; calc. for C<sub>14</sub>H<sub>21</sub>SiO<sub>3</sub>: 265.1260.

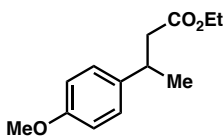
### Ethyl 2-((4-(methoxymethyl)phenyl)dimethylsilyl)acetate (**52h'**)

This compound was prepared by rhodium-catalyzed Si–H insertion. To a dry 50 mL round bottom flask, under argon, was added (4-(methoxymethyl)phenyl)dimethylsilane (**51h**) (541 mg, 3 mmol, 1.0 equiv.),  $\text{Rh}_2(\text{OAc})_4$  (13.3 mg, ~1 mol%), and anhydrous dichloromethane (12 mL). The mixture was cooled to  $-78\text{ }^\circ\text{C}$ , after which ethyl diazoacetate (393 mg, 3.0 mmol, 1.0 equiv.) in dichloromethane (3 mL) was added dropwise to the solution over 2 hours. The reaction was allowed to slowly warm to room temperature and stirred for a total of 12 hours. The crude reaction mixture was filtered through a pad of Celite and concentrated under reduced pressure. The crude mixture was purified by silica column chromatography using hexanes / ethyl acetate to deliver **52h'** with impurities. A second purification by silica column chromatography using hexanes / diethyl ether/ dichloromethane afforded **52h'** (92.6 mg, 0.35 mmol, 12% yield).



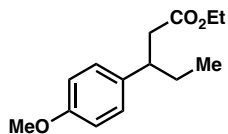
$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.52 (d,  $J = 8.1$  Hz, 2H), 7.34 (d,  $J = 8.1$  Hz, 2H), 4.46 (s, 2H), 4.04 (q,  $J = 7.2$  Hz, 2H), 3.39 (s, 3H), 2.11 (s, 2H), 1.16 (t,  $J = 7.1$  Hz, 3H), 0.40 (s, 6H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  172.7, 139.7, 136.4, 133.8, 127.2, 74.6, 60.1, 58.3, 26.4, 14.5, -2.6. HRMS (FAB)  $m/z$ : 265.1260  $[(M + H^+) - \text{H}_2]$ ; calc. for  $\text{C}_{14}\text{H}_{21}\text{SiO}_3$ : 265.1260.

### Ethyl 3-(4-methoxyphenyl)butanoate (**52i**)



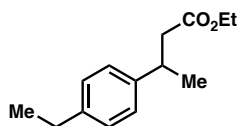
This compound was prepared from 1-(4-methoxyphenyl)ethan-1-one using **General Procedure C**.  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  7.14 (d,  $J = 8.5$  Hz, 2H), 6.84 (d,  $J = 8.7$  Hz, 2H), 4.08 (qd,  $J = 7.2, 1.2$  Hz, 2H), 3.79 (s, 3H), 3.24 (h,  $J = 7.1$  Hz, 1H), 2.57 (dd,  $J = 14.9, 7.2$  Hz, 1H), 2.51 (dd,  $J = 14.9, 8.0$  Hz, 1H), 1.28 (d,  $J = 7.0$  Hz, 3H), 1.19 (t,  $J = 7.1$  Hz, 3H). Spectral data are in agreement with that for the enzymatic product (see Section C.8).

### Ethyl 3-(4-methoxyphenyl)pentanoate (**52j**)

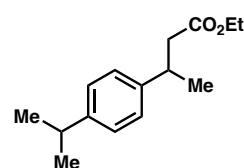


This compound was prepared from 1-(4-methoxyphenyl)propan-1-one using **General Procedure C**. Spectral data are in agreement with literature report<sup>28</sup>.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.09 (d,  $J = 8.6$  Hz, 2H), 6.83 (d,  $J = 8.8$  Hz, 2H), 4.03 (qd,  $J = 7.2, 1.3$  Hz, 2H), 3.78 (s, 3H), 2.95 (tdd,  $J = 9.0, 7.0, 5.3$  Hz, 1H), 2.60 (dd,  $J = 15.0, 7.0$  Hz, 1H), 2.51 (dd,  $J = 14.9, 8.3$  Hz, 1H), 1.68 (ddq,  $J = 13.3, 7.4, 5.4$  Hz, 1H), 1.56 (ddq,  $J = 13.5, 9.4, 7.3$  Hz, 1H), 1.14 (t,  $J = 7.1$  Hz, 3H), 0.78 (t,  $J = 7.4$  Hz, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  172.7, 158.2, 136.1, 128.5, 113.8, 60.3, 55.3, 43.3, 41.9, 29.4, 14.3, 12.1.



**Ethyl 3-(4-ethylphenyl)butanoate (52k)**

This compound was prepared from 1-(4-ethylphenyl)ethan-1-one using **General Procedure C**.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.14 (*app.* s, 4H), 4.08 (q,  $J$  = 7.1 Hz, 2H), 3.25 (dp,  $J$  = 8.3, 7.0 Hz, 1H), 2.66 – 2.48 (m, 4H), 1.29 (d,  $J$  = 6.9 Hz, 3H), 1.26 – 1.15 (m, 6H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  172.7, 143.1, 142.3, 128.0, 126.8, 60.4, 43.2, 36.2, 28.5, 22.0, 15.7, 14.3. **HRMS** (FAB)  $m/z$ : 221.1532 ( $\text{M} + \text{H}^+$ ); calc. for  $\text{C}_{14}\text{H}_{21}\text{O}_2$ : 221.1542

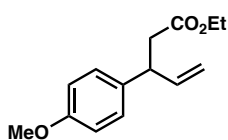
**Ethyl 3-(4-isopropylphenyl)butanoate (52l)**

This compound was prepared from 1-(4-isopropylphenyl)ethan-1-one using **General Procedure C**.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.15 (*app.* s, 4H), 4.08 (q,  $J$  = 7.1 Hz, 2H), 3.25 (dp,  $J$  = 8.5, 6.9 Hz, 1H), 2.87 (hept,  $J$  = 6.9 Hz, 1H), 2.60 (dd,  $J$  = 14.9, 6.7 Hz, 1H), 2.51 (dd,  $J$  = 14.9, 8.4 Hz, 1H), 1.29 (d,  $J$  = 7.0 Hz, 3H), 1.23 (d,  $J$  = 6.9 Hz, 6H), 1.18 (t,  $J$  = 7.1 Hz, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  172.7, 147.0, 143.2, 126.8, 126.6, 60.4, 43.3, 36.2, 33.8, 24.2, 21.9, 14.3. **HRMS** (FAB)  $m/z$ : 235.1696 ( $\text{M} + \text{H}^+$ ); calc. for  $\text{C}_{15}\text{H}_{23}\text{O}_2$ : 235.1698.

**Ethyl 3-(4-methoxyphenyl)pent-4-enoate (52m)**

This compound was accessed in a two-step sequence. First, *p*-methoxycinnamaldehyde (811 mg, 5 mmol, 1.0 equiv.) was reduced using  $\text{NaBH}_4$  (227 mg, 6 mmol, 1.2 equiv.) in methanol (15 mL) under standard reaction conditions (0 °C for 2 hours). The reaction mixture was quenched with  $\text{NH}_4\text{Cl}$  (sat. aq., 10 mL) and diluted with dichloromethane (15 mL). Phases were separated and the aqueous layer was extracted with dichloromethane (4  $\times$  15 mL). The combined organics were washed with brine (25 mL), dried over  $\text{Na}_2\text{SO}_4$ , and concentrated under reduced pressure. Purification by silica column chromatography with hexanes / ethyl acetate delivered *p*-methoxycinnamyl alcohol (752 mg, 4.6 mmol, 92% yield), with spectral data that match literature report<sup>29</sup>.

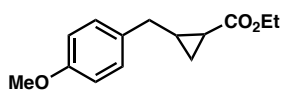
Next, to a 50 mL round bottom flask equipped with short-path condenser were added *p*-methoxycinnamyl alcohol (740 mg, 4.5 mmol, 1.0 equiv.), triethyl orthoacetate (7.3 g, 45 mmol, 10 equiv.), and propionic acid (52 mg, 0.7 mmol, 0.15 equiv.). Following standard Johnson-Claisen rearrangement conditions, this mixture was heated to 140 °C until complete conversion of *p*-methoxycinnamyl alcohol was observed by TLC (~23 hours). Additional propionic acid (2  $\times$  52 mg) was added after 6 hours and 9 hours reaction time. The reaction mixture was removed from heat, concentrated under reduced pressure, and purified using silica gel chromatography with hexanes / ethyl acetate as eluents. A second purification by silica gel chromatography with hexanes / ether afforded **52m** (357 mg, 1.6 mmol, 36% yield).



Spectral data for **52m** are in agreement with literature report<sup>30</sup>. **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$  7.13 (d,  $J$  = 8.6 Hz, 2H), 6.85 (d,  $J$  = 8.8 Hz, 2H), 5.96 (ddd,  $J$  = 17.5, 9.9, 6.9 Hz, 1H), 5.09 – 5.05 (m, 1H), 5.03 (dt,  $J$  = 5.4, 1.3 Hz, 1H), 4.07 (qd,  $J$  = 7.1, 1.0 Hz, 2H), 3.86 – 3.80 (m, 1H), 3.78 (s, 3H), 2.73 (dd,  $J$  = 15.0, 8.0 Hz, 1H), 2.65 (dd,  $J$  = 15.0, 7.6 Hz, 1H), 1.18 (t,  $J$  = 7.1 Hz, 3H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>)  $\delta$  172.1, 158.4, 140.7, 134.6, 128.6, 114.6, 114.0, 60.5, 55.4, 44.9, 40.6, 14.3.

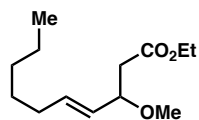
### Ethyl 2-(4-methoxybenzyl)cyclopropane-1-carboxylate (**52m'**)

This compound was prepared by rhodium-catalyzed alkene cyclopropanation. To a dry 100 mL round bottom flask, under argon, were added 4-allylanisole (3.0 g, 20 mmol, 10 equiv.), Rh<sub>2</sub>(OAc)<sub>4</sub> (8.8 mg, ~1 mol%), and anhydrous dichloromethane (10 mL). Ethyl diazoacetate (262 mg, 2 mmol, 1.0 equiv.) in dichloromethane (10 mL) was added over ~8 hours using a syringe pump; the reaction mixture was allowed to stir for a total of 20 hours at room temperature. The reaction mixture was diluted with diethyl ether (20 mL), filtered through a pad of Celite, and concentrated under reduced pressure. Several rounds of purification by silica column chromatography with hexanes / ethyl acetate or hexanes / diethyl ether eluent systems afforded *cis*-**52m'** and *trans*-**52m'** as individual isomers (combined mass 148.1 mg, 0.632 mmol, 32% yield).

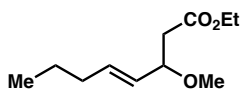


Spectral data are in agreement with literature report<sup>31</sup>. Characterization data for *cis*-**52m'**: **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$  7.13 (d,  $J$  = 8.7 Hz, 2H), 6.83 (d,  $J$  = 8.8 Hz, 2H), 4.13 (q,  $J$  = 7.2 Hz, 2H), 3.79 (s, 3H), 2.86 (dd,  $J$  = 14.9, 6.9 Hz, 1H), 2.77 (dd,  $J$  = 15.0, 7.6 Hz, 1H), 1.77 (ddd,  $J$  = 8.8, 7.6, 5.9 Hz, 1H), 1.56 – 1.44 (m, 1H), 1.24 (t,  $J$  = 7.1 Hz, 3H), 1.14 – 1.06 (m, 2H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>)  $\delta$  173.1, 158.0, 133.7, 129.3, 113.9, 60.5, 55.4, 32.1, 23.1, 18.7, 14.5, 13.7. Characterization data for *trans*-**52m'**: **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$  7.12 (d,  $J$  = 8.7 Hz, 2H), 6.84 (d,  $J$  = 8.7 Hz, 2H), 4.11 (qd,  $J$  = 7.1, 1.1 Hz, 2H), 3.79 (s, 3H), 2.71 (dd,  $J$  = 14.7, 6.3 Hz, 1H), 2.52 (dd,  $J$  = 14.8, 7.1 Hz, 1H), 1.65 (ddtd,  $J$  = 8.7, 7.1, 6.4, 4.1 Hz, 1H), 1.52 – 1.46 (m, 1H), 1.27 – 1.20 (m, 4H), 0.81 (ddd,  $J$  = 8.2, 6.3, 4.2 Hz, 1H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>)  $\delta$  174.4, 158.2, 132.3, 129.5, 113.9, 60.5, 55.4, 37.6, 23.4, 20.3, 15.3, 14.4.

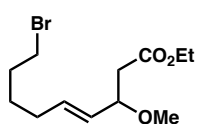
### Ethyl (*E*)-3-methoxydec-4-enoate (**54a**)



This compound was prepared from (*E*)-oct-2-enal using **General Procedure B-2**. **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$  5.69 (dt,  $J$  = 15.4, 6.8 Hz, 1H), 5.28 (ddt,  $J$  = 15.4, 8.3, 1.5 Hz, 1H), 4.14 (qd,  $J$  = 7.2, 0.8 Hz, 2H), 3.97 (td,  $J$  = 8.2, 5.5 Hz, 1H), 3.25 (s, 3H), 2.59 (dd,  $J$  = 14.9, 8.1 Hz, 1H), 2.42 (dd,  $J$  = 14.9, 5.5 Hz, 1H), 2.10 – 1.97 (m, 2H), 1.43 – 1.20 (m, 9H), 0.88 (t,  $J$  = 6.9 Hz, 3H). Spectral data are in agreement with that for the enzymatic product (see Section C.8).

**Ethyl (*E*)-3-methoxyoct-4-enoate (54b)**

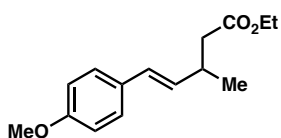
This compound was prepared from (*E*)-hex-2-enal using **General Procedure B-2**.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  5.69 (dt,  $J = 15.4$ , 6.8 Hz, 1H), 5.29 (ddt,  $J = 15.4$ , 8.2, 1.5 Hz, 1H), 4.14 (qd,  $J = 7.1$ , 0.8 Hz, 2H), 3.97 (td,  $J = 8.1$ , 5.5 Hz, 1H), 3.25 (s, 3H), 2.59 (dd,  $J = 14.9$ , 8.1 Hz, 1H), 2.42 (dd,  $J = 14.9$ , 5.6 Hz, 1H), 2.06 – 1.99 (m, 2H), 1.40 (sext,  $J = 7.3$  Hz, 2H), 1.25 (t,  $J = 7.1$  Hz, 3H), 0.89 (t,  $J = 7.4$  Hz, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  171.2, 135.3, 128.9, 79.0, 60.6, 56.2, 41.5, 34.3, 22.4, 14.4, 13.7. **HRMS** (FAB)  $m/z$ : 199.1320 [ $(\text{M} + \text{H}^+) - \text{H}_2$ ]; calc. for  $\text{C}_{11}\text{H}_{19}\text{O}_3$ : 199.1334.

**Ethyl (*E*)-9-bromo-3-methoxynon-4-enoate (54c)**

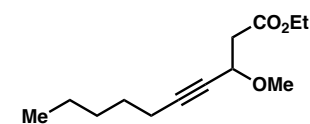
This compound was prepared from (*E*)-7-bromohept-2-enal using **General Procedure B-2**. The synthesis of (*E*)-7-bromohept-2-enal was described in the synthesis of compound **53c** in Section C.5. This compound was prepared by Dr. Xiongyi Huang.  $^1\text{H}$  NMR (400 MHz, Chloroform-*d*)  $\delta$  5.67 (dt,  $J = 15.4$ , 6.7 Hz, 1H), 5.31 (dd,  $J = 15.4$ , 8.1 Hz, 1H), 4.13 (q,  $J = 7.1$  Hz, 2H), 3.97 (td,  $J = 8.0$ , 5.5 Hz, 1H), 3.40 (t,  $J = 6.7$  Hz, 2H), 3.25 (s, 3H), 2.59 (dd,  $J = 15.0$ , 8.0 Hz, 1H), 2.41 (dd,  $J = 15.0$ , 5.6 Hz, 1H), 2.08 (q,  $J = 7.2$  Hz, 2H), 1.91 – 1.79 (m, 2H), 1.53 (p,  $J = 7.5$  Hz, 2H), 1.25 (t,  $J = 7.2$  Hz, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  171.1, 134.3, 129.5, 78.8, 60.6, 56.3, 41.4, 33.7, 32.2, 31.4, 27.7, 14.4. **HRMS** (FAB)  $m/z$ : 293.0764 ( $\text{M} + \text{H}^+$ ); calc. for  $\text{C}_{12}\text{H}_{22}\text{O}_3^{79}\text{Br}$ : 293.0752.

**Ethyl (*E*)-5-(4-methoxyphenyl)-3-methylpent-4-enoate (54d)**

To a 6 mL vial equipped with a stir bar was added Grubbs' catalyst 2<sup>nd</sup> generation (10 mg, 2 mol%). The vial was then evacuated and backfilled with argon for three times. Under argon, a dry  $\text{CH}_2\text{Cl}_2$  solution (2 mL) containing 4-vinylanisole (100 mg, 0.75 mmol) and ethyl 3-methylpent-4-enoate (503 mg, 3.75 mmol) was added to the vial via syringe. The mixture was stirred at 40 °C for 24 hours and then cooled to room temperature and filtered through a silica plug. The solvent was removed under reduced pressure and the crude product was purified using silica column chromatography with hexanes / ethyl acetate to give **54d** (37 mg, 20% yield).



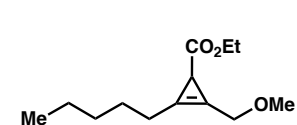
$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.30 – 7.24 (m, 2H), 6.84 (d,  $J = 8.8$  Hz, 2H), 6.34 (d,  $J = 15.9$  Hz, 1H), 5.99 (dd,  $J = 15.9$ , 7.6 Hz, 1H), 4.12 (q,  $J = 7.1$  Hz, 2H), 3.80 (s, 3H), 2.90 – 2.75 (m, 1H), 2.41 (dd,  $J = 14.7$ , 7.3 Hz, 1H), 2.34 (dd,  $J = 14.7$ , 7.3 Hz, 1H), 1.23 (t,  $J = 7.1$  Hz, 3H), 1.14 (d,  $J = 6.7$  Hz, 3H). Spectral data are in agreement with that for the enzymatic product (see Section C.8).

**Ethyl 3-methoxydec-4-ynoate (54e)**

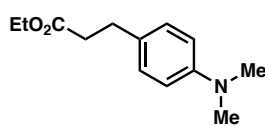
This compound was prepared from oct-2-ynal using **General Procedure B-2**. This compound was prepared by Kai Chen. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.39 (ddt, *J* = 8.3, 5.4, 2.0 Hz, 1H), 4.16 (qd, *J* = 7.2, 1.0 Hz, 2H), 3.39 (s, 3H), 2.73 (dd, *J* = 15.5, 8.4 Hz, 1H), 2.63 (dd, *J* = 15.5, 5.4 Hz, 1H), 2.20 (td, *J* = 7.1, 2.0 Hz, 2H), 1.50 (p, *J* = 7.2 Hz, 2H), 1.41 – 1.29 (m, 4H), 1.26 (t, *J* = 7.1 Hz, 3H), 0.89 (t, *J* = 7.1 Hz, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 170.4, 87.3, 67.7, 60.8, 56.6, 41.7, 31.1, 28.4, 22.3, 18.8, 14.3, 14.1 (one carbon may be overlapping with the solvent peaks). HRMS (FAB) *m/z*: 227.1638 (*M* + *H*<sup>+</sup>); calc. for C<sub>13</sub>H<sub>23</sub>O<sub>3</sub>: 227.1647.

**Ethyl 2-(methoxymethyl)-3-pentylcycloprop-2-ene-1-carboxylate (54e')**

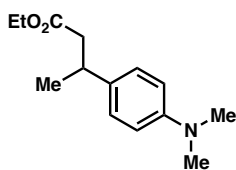
This compound was prepared by rhodium-catalyzed cyclopropanation. To a dry 50 mL round bottom flask was added 1-methoxyoct-2-yne (**54e**) (280 mg, 2.0 mmol, 1.0 equiv.), Rh<sub>2</sub>(OAc)<sub>4</sub> (9.0 mg, 1 mol%), and anhydrous dichloromethane (6 mL). The mixture was cooled to -78 °C, after which ethyl diazoacetate (87%, 525 mg, 4.0 mmol, 2.0 equiv.) in dichloromethane (5 mL) was added dropwise to the solution over 6 hours. The reaction was allowed to slowly warm to room temperature and stirred for a total of 18 hours. The reaction mixture was concentrated under reduced pressure. The crude product was purified by silica column chromatography using hexanes / ethyl acetate, followed by C18 column using methanol / water, to afford **54e'** (26 mg, 0.11 mmol, 6% yield). This compound was prepared by Kai Chen.



<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.37 (t, *J* = 1.6 Hz, 2H), 4.12 (q, *J* = 7.1 Hz, 2H), 3.39 (s, 3H), 2.47 (tt, *J* = 7.5, 1.6 Hz, 2H), 2.20 (s, 1H), 1.64 – 1.52 (m, 2H), 1.37 – 1.28 (m, 4H), 1.24 (t, *J* = 7.1 Hz, 3H), 0.93 – 0.86 (m, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 176.2, 110.3, 102.5, 65.8, 60.2, 58.6, 31.5, 26.7, 24.7, 22.7, 22.5, 14.5, 14.1. HRMS (EI) *m/z*: 226.1573 (*M*<sup>+</sup>); calc. for C<sub>13</sub>H<sub>22</sub>O<sub>3</sub>: 226.1569.

**Ethyl 3-(4-(dimethylamino)phenyl)propanoate (57a')**

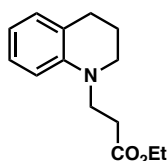
This compound was prepared from 4-(dimethylamino)benzaldehyde using **General Procedure C**. This compound was prepared by Kai Chen. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.09 (d, *J* = 8.7 Hz, 2H), 6.70 (d, *J* = 8.7 Hz, 2H), 4.13 (q, *J* = 7.2 Hz, 2H), 2.92 (s, 6H), 2.89 – 2.82 (m, 2H), 2.61 – 2.54 (m, 2H), 1.25 (t, *J* = 7.1 Hz, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 173.4, 149.4, 129.0, 128.8, 113.1, 60.4, 41.0, 36.6, 30.2, 14.4. HRMS (EI) *m/z*: 221.1430 (*M*<sup>+</sup>); calc. for C<sub>13</sub>H<sub>19</sub>NO<sub>2</sub>: 221.1416.

**Ethyl 3-(4-(dimethylamino)phenyl)butanoate (57b')**

This compound was prepared from 1-(4-(dimethylamino)phenyl)ethan-1-one using **General Procedure C**. Spectral data are in agreement with literature report<sup>32</sup>. This compound was prepared by Kai Chen. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.10 (d, *J* = 8.7 Hz, 2H), 6.70 (d, *J* = 8.7 Hz, 2H), 4.08 (qd, *J* = 7.1, 1.1 Hz, 2H), 3.20 (dt, *J* = 8.4, 6.8 Hz, 1H), 2.92 (s, 6H), 2.57 (dd, *J* = 14.8, 6.8 Hz, 1H), 2.49 (dd, *J* = 14.8, 8.4 Hz, 1H), 1.27 (d, *J* = 7.0 Hz, 3H), 1.20 (t, *J* = 7.1 Hz, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 172.8, 149.4, 134.0, 127.4, 113.0, 60.3, 43.5, 41.0, 35.7, 22.0, 14.4.

**Ethyl 3-(3,4-dihydroquinolin-1(2*H*)-yl)propanoate (57f')**

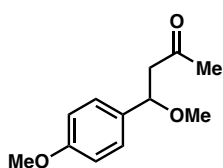
To a 100-mL round-bottom flask were added 1,2,3,4-tetrahydroquinoline (266.4 mg, 2.0 mmol, 1.0 equiv.), ethyl 3-bromopropanoate (0.97 mL, 6.0 mmol, 3.0 equiv.), K<sub>2</sub>CO<sub>3</sub> (0.552 g, 4.0 mmol, 2.0 equiv.), KI (66.0 mg, 0.4 mmol, 0.2 equiv.) and *N,N*-dimethylformamide (30 mL). The reaction mixture was heated at 120 °C for 4 hours. After the reaction was cooled to room temperature and quenched by H<sub>2</sub>O (40 mL), the crude product was extracted by diethyl ether (20 mL × 3). The combined organic layer was washed by H<sub>2</sub>O (40 mL) and brine (40 mL), and then dried over sodium sulfate and concentrated under reduced pressure. The crude product was purified by silica column chromatography with pentane / diethyl ether, followed by C18 column with methanol / water, to afford **57f'** (350 mg, 1.5 mmol, 75% yield). This compound was prepared by Kai Chen.



This compound is known in the literature<sup>33</sup>. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.11 – 7.01 (m, 1H), 6.95 (dq, *J* = 7.1, 1.1 Hz, 1H), 6.66 – 6.54 (m, 2H), 4.14 (q, *J* = 7.1 Hz, 2H), 3.65 – 3.57 (m, 2H), 3.33 – 3.25 (m, 2H), 2.75 (t, *J* = 6.4 Hz, 2H), 2.64 – 2.54 (m, 2H), 1.99 – 1.89 (m, 2H), 1.26 (t, *J* = 7.1 Hz, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 172.6, 144.6, 129.5, 127.3, 122.8, 116.2, 110.7, 60.7, 49.5, 47.4, 31.5, 28.1, 22.3, 14.4.

**4-Methoxy-4-(4-methoxyphenyl)butan-2-one (59d)**

This compound was prepared according to the procedure of Yadav *et al*<sup>34</sup>. Briefly, a mixture of 4-anisaldehyde (10 mmol), 2,2-dimethoxypropane (20 mmol) and iodine (0.2 mmol) in dry methylene chloride (20 mmol) was stirred under N<sub>2</sub> for 30 min. After the reaction was complete as indicated by TLC, the reaction mixture was diluted with water and extracted with ethyl acetate (2 × 30 mL). The combined organic extracts were washed with sodium thiosulfate (aq., 15% w/v) and brine, and then dried over sodium sulfate and concentrated under reduced pressure. The crude product was purified by silica column chromatography with hexanes / ethyl acetate. This compound was prepared by Dr. Xiongyi Huang.



$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.25 – 7.21 (m, 2H), 6.88 (d,  $J$  = 8.8 Hz, 2H), 4.58 (dd,  $J$  = 8.8, 4.5 Hz, 1H), 3.79 (s, 3H), 3.16 (s, 3H), 3.05 – 2.88 (m, 1H), 2.57 (dd,  $J$  = 15.8, 4.5 Hz, 1H), 2.14 (s, 3H). Spectral data are in agreement with that for the enzymatic product (see Section C.8).

### C.7 Small scale enzymatic reactions and product calibration curves

Enzymatic reactions performed on analytical scale were conducted following the general procedure described in Chapter 4.5 Experimental Methods. Product formation was quantified by HPLC or GC based on the calibration curve of the corresponding racemic reference compound. TTN is defined as the amount of product divided by total heme protein as measured by the hemochrome assay. Calibration curves and data analysis for results shown in Figures 4-6 and 4-7 are available in the Supplementary Information of the published paper.

#### HPLC calibration curve preparation

Stock solutions of chemically synthesized products at various concentrations (1 to 200 mM in EtOH) were prepared. To a 2 mL vial were added 380  $\mu\text{L}$  water, 20  $\mu\text{L}$  product stock solution, 10  $\mu\text{L}$  internal standard (60 mM ethyl phenoxyacetate or 40 mM ethyl benzoate in acetonitrile, as appropriate) and 400  $\mu\text{L}$  acetonitrile. The mixture was vortexed and analyzed by HPLC. Data points represent the average of two runs. The standard curves plot product concentration in mM (y-axis) against the ratio of product area to internal standard area on the HPLC (x-axis).

#### GC calibration curve preparation

Stock solutions of chemically synthesized products at various concentrations (0.2 to 200 mM in EtOH) were prepared. To a microcentrifuge tube were added 380  $\mu\text{L}$  M9-N buffer, 20  $\mu\text{L}$  product stock solution, 10  $\mu\text{L}$  internal standard (40 mM 1,3,5-trimethoxybenzene in cyclohexane), and 800  $\mu\text{L}$  mixed solvent system (cyclohexane : ethyl acetate = 1:1). The mixture was vortexed (10 seconds, 3 times) then centrifuged ( $20,000 \times g$ , 5 min) to completely separate the organic and aqueous layers. The organic layer was removed for GC analysis. All data points represent the average of at least two runs. The standard curves plot product concentration in mM (y-axis) against the ratio of product area to internal standard area on the GC (x-axis).

## C.8 Enzymatic C–H alkylation reactions on preparative scale

### Preparation of *E. coli* cells expressing P411 variants for preparative-scale reactions

HB<sub>amp</sub> (480 mL) in a 1 L flask was inoculated with an overnight culture (20 mL, LB<sub>amp</sub>) of *E. coli* (*E. coli* BL21(DE3)) cells containing a pET22b(+) plasmid encoding the desired P411 variant. The culture was shaken at 37 °C and 220 rpm (no humidity control) for 2.5 hours. The culture was placed on ice for 40 minutes, and 5-aminolevulinic acid (1.0 mM final concentration) and IPTG (0.5 mM final concentration) were added. The incubator temperature was reduced to 20 °C, and the culture was allowed to shake for 16–18 hours at 130 rpm. When greater amounts of cells were required, additional *E. coli* cultures were prepared in this manner. Cells were pelleted by centrifugation (3,000 × g, 5 min, 4 °C), resuspended in M9-N buffer and adjusted to OD<sub>600</sub> ≈ 60. An aliquot of cells at OD<sub>600</sub> ≈ 60 (3 mL) was taken for the hemochrome assay to determine protein concentration. When applicable, the cell suspension was diluted with M9-N buffer to achieve the OD<sub>600</sub> which was used for the reaction. Cell suspensions in M9-N buffer were kept on ice until use.

### General Procedure D: Enzymatic C–H alkylation reactions on preparative scale

*Procedure D-I (alkane substrate is limiting reagent).* To an Erlenmeyer flask equipped with a screw cap (reaction vessel, Chemglass CG-1543: 250 mL, 500 mL, or 1000 mL) was added a suspension of *E. coli* expressing the indicated P411 variant (generally OD<sub>600</sub> = 30). The headspace of the reaction vessel was degassed with argon (at least 1 hour for volumes less than 200 mL, at least 2 hours for volumes greater than 200 mL) while kept on ice. To degas the headspace of a flask containing *E. coli* cells, the flask is covered with aluminum foil and a stream of argon is flowed through the flask just above the cell suspension. Separately, a solution of *D*-glucose (250 mM in M9-N) was bubbled with argon and the headspace of a flask containing GOX oxygen depletion system (a solution of 14,000 U/mL catalase and 1,000 U/mL glucose oxidase, kept on ice) was degassed for at least 1 hour. In an anaerobic chamber, GOX, *D*-glucose, alkane substrate (1.0 equiv.), and diazo compound (1.0 equiv.) were added to the reaction vessel in this order. The vessel was capped, sealed with parafilm, and shaken (150–200 rpm) at room temperature. After one hour, the reaction vessel was transferred again to the anaerobic chamber where a second portion of *E. coli* cells expressing the P411 variant (the headspace of the flask containing these cells was degassed with argon following same procedure) and additional diazo compound (1.0 equiv.) were added. The vessel was capped, sealed with parafilm, and shaken (150–200 rpm) at room temperature for 14–17 additional hours. Final conditions were *E. coli* expressing P411 variant, alkane substrate (0.2–0.5 mmol, 1.0 equiv., larger scales for **52a**, **54a**, and **57f**, see Section C.9), diazo compound (2.0 equiv.), *D*-glucose (25 mM), GOX oxygen depletion system (700 U/mL catalase, 50 U/mL glucose oxidase), 2 vol% EtOH in M9-N buffer under anaerobic conditions; total reaction time 15–18 hours.

*Procedure D-II (diazo compound is limiting reagent).* To an Erlenmeyer flask

equipped with a screw cap (reaction vessel, Chemglass CG-1543: 250 mL, 500 mL, or 1000 mL) was added a suspension of *E. coli* expressing the indicated P411 variant. The headspace of the reaction vessel was degassed with argon (at least 1 hour) while kept on ice. To degas the headspace of a flask containing *E. coli* cells, the flask is covered with aluminum foil and a stream of argon is flowed through the flask just above the cell suspension. Separately, a solution of *D*-glucose (250 mM in M9-N) was bubbled with argon. In an anaerobic chamber, *D*-glucose, alkane substrate (2.0 equiv.), and diazo compound (1.0 equiv.) were added to the reaction vessel in this order. The vessel was capped, sealed with parafilm, and shaken (150–200 rpm) at room temperature. Final conditions were *E. coli* expressing P411 variant, alkane substrate (1.0 mmol, 2.0 equiv.), diazo compound (0.5 mmol, 1.0 equiv.), *D*-glucose (25 mM), 2 vol% EtOH in M9-N buffer under anaerobic conditions; total reaction time 15–18 hours.

*Workup Procedure D-i.* Every 35 mL portion of the preparative scale reaction mixture was transferred to a 50 mL Eppendorf conical tube (catalog no. 0030122178). To the reaction mixture in every tube was added 15 mL mixed organic solvent (1 : 1 hexanes : ethyl acetate); the solution was shaken vigorously and centrifuged ( $10,000 \times g$ , 5 minutes, Beckman-Coulter Avanti J-25 centrifuge equipped with JA-12 rotor) to separate the organic and aqueous layers. The organic layer was collected and the aqueous layer was subject to three additional rounds of extraction. The organic layers were combined, dried over  $\text{Na}_2\text{SO}_4$ , and concentrated under reduced pressure. Purification was performed by silica column chromatography with either hexanes / ethyl acetate or hexanes / dichloromethane / diethyl ether as eluent systems to afford the desired product. Additional purification by reverse phase chromatography (Biotage Isolera equipped with Biotage SNAP Ultra C18 column, water / methanol eluent system) was utilized if necessary. TTNs were calculated based on measured protein concentration and the isolated yield of the product.

*Workup Procedure D-ii.* Every ~100–125 mL portion of preparative scale reaction mixture was transferred to a centrifuge bottle. To the reaction mixture in every bottle was added equal volume ethyl acetate; the solution was shaken vigorously and centrifuged ( $14,000 \times g$ , 10 minutes, Beckman-Coulter Avanti J-25 centrifuge equipped with JA-10 rotor) to separate the organic and aqueous layers. The organic layer was collected and the aqueous layer was subject to three additional rounds of extraction. The combined organics were dried over  $\text{Na}_2\text{SO}_4$ , and concentrated under reduced pressure. Purification was performed by silica column chromatography with either hexanes / ethyl acetate or hexanes / dichloromethane / diethyl ether as eluent systems to afford the desired product. Additional purification by reverse phase chromatography (Biotage Isolera equipped with Biotage SNAP Ultra C18 column, water / methanol eluent system) was utilized if necessary. TTNs were calculated based on measured protein concentration and the isolated yield of the product.



### Ethyl 3-methoxy-3-(4-methoxyphenyl)propanoate (52a)

This compound was prepared using a modified version of **General Procedure D-I** carried out at 4 °C. The reaction mixture was kept on ice during the addition of all reagents and shaken in an incubator set to 4 °C and 150 rpm. In addition, a solution of Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub> (5 mL, 40 mM in M9-N, bubbled with argon for 1 hour) was added following the second addition of *E. coli* cells expressing P411-CHF. Final conditions were *E. coli* expressing P411-CHF (resuspended to OD<sub>600</sub> = 44), alkane substrate (1.0 mmol, 1.0 equiv.), ethyl diazoacetate (2.0 mmol, 2.0 equiv.), *D*-glucose (25 mM), GOX oxygen depletion system (700 U/mL catalase, 50 U/mL glucose oxidase), Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub> (1 mM), 2 vol% EtOH in M9-N buffer under anaerobic conditions at 4 °C; total reaction time 18 hours.

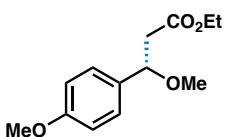
The reaction was quenched with the addition of 100 mL acetonitrile. The crude reaction mixture was transferred to a centrifuge tube, shaken vigorously, and centrifuged to pellet the cells (14,000 × g, 10 minutes). The supernatant was decanted and the acetonitrile was removed under reduced pressure. Following, the aqueous layer was extracted with ethyl acetate (3 × 100 mL). The cell pellet was resuspended in H<sub>2</sub>O and this suspension was also extracted using ethyl acetate (3 × 20 mL; centrifugation (3,000 × g, 5 min) was used to help separate the organic and aqueous layers. The combined organics were dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. Purification by reverse phase preparative HPLC (column: Eclipse XDB-C8, 5 µm, 9.4 × 250 mm) using water / acetonitrile eluent system afforded **52a**.

<i>E. coli</i> suspension in M9-N (variant: P411-CHF, OD <sub>600</sub> = 44, 4 °C)				GOX solution <sup>a</sup>	<i>D</i> -glucose in M9-N <sup>b</sup>
Addition 1, volume/ mL	Addition 2, volume/mL	[PC]/µM	n_pro/µmol	volume/mL	volume/mL
110.0	51.0 <sup>*</sup>	4.77	0.768	10.0	20.0
Alkane substrate (51a) stock in EtOH			Ethyl diazoacetate stock in EtOH		
stock/M	volume/mL	n_1/mmol	stock/M	volume/mL add. 1; add. 2	n_2/mmol
0.50	2.0	1.0	1.0	1.0; 1.0	2.0
Purification eluent		Product			
1: water / MeCN (reverse phase preparative HPLC)		m[Pdt]/mg	n[Pdt]/mmol	yield	TTN
		194.7	0.817	82%	1060

<sup>\*</sup>A solution of Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub> (5 mL, 40 mM in M9-N, bubbled with argon for 1 hour) was added following the second addition of *E. coli* cells expressing P411-CHF. Final concentration of Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub> in the reaction is 1 mM.

**Notes:** [PC] = protein concentration in original cell suspension, n\_pro = amount of protein in the reaction, n\_1 = amount of alkane substrate in the reaction, n\_2 = total amount of diazo compound in the reaction, add. = addition, m[Pdt] = mass of product isolated, n[Pdt] = amount of product. <sup>a</sup>GOX oxygen depletion system is 14,000 U/mL catalase and 1,000 U/ mL glucose oxidase in 0.1 M potassium phosphate buffer (pH=8.0); the final reaction mixture contains 700 U/mL catalase and 50 U/mL glucose oxidase. <sup>b</sup>*D*-glucose stock

solution is 250 mM in M9-N buffer; final concentration of *D*-glucose in the reaction is 25 mM. These notes apply for all tables in this section.

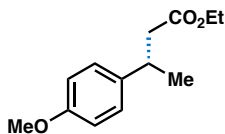


**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.25 (d, *J* = 8.5 Hz, 2H), 6.89 (d, *J* = 8.8 Hz, 2H), 4.58 (dd, *J* = 9.0, 4.9 Hz, 1H), 4.14 (qd, *J* = 7.1, 1.2 Hz, 2H), 3.81 (s, 3H), 3.19 (s, 3H), 2.80 (dd, *J* = 15.2, 9.0 Hz, 1H), 2.55 (dd, *J* = 15.2, 4.9 Hz, 1H), 1.23 (t, *J* = 7.1 Hz, 3H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 171.2, 159.5, 132.7, 128.0, 114.0, 79.7, 60.7, 56.7, 55.4, 43.7, 14.3. **HRMS** (EI) *m/z*: 238.1213 (*M*<sup>+</sup>); calc. for C<sub>13</sub>H<sub>18</sub>O<sub>4</sub>: 238.1205. [ $\alpha$ ]<sub>D</sub><sup>23</sup> = -46.354 ± 0.411° (*c* 0.5, CHCl<sub>3</sub>). **SFC Chiralpak AD-H column** (3% *i*-PrOH in supercritical CO<sub>2</sub>, 2.5 mL/min, 40 °C), *t<sub>r</sub>* = 9.02 min (major), 10.50 min (minor), 98.0 : 2.0 e.r.

### Ethyl 3-(4-methoxyphenyl)butanoate (52i)

This compound was prepared using **General Procedure D-I** and Workup Procedure D-*i*.

<i>E. coli</i> suspension in M9-N (variant: P411-CHF, OD <sub>600</sub> = 29)				GOX solution <sup>a</sup>	<i>D</i> -glucose in M9-N <sup>b</sup>
Addition 1, volume/ mL	Addition 2, volume/mL	[PC]/μM	n_pro/μmol	volume/mL	volume/mL
110.0	56.0	2.25	0.374	10.0	20.0
Alkane substrate (51i) stock in EtOH			Ethyl diazoacetate stock in EtOH		
stock/M	volume/mL	n_1/mmol	stock/M	volume/mL add. 1; add. 2	n_2/mmol
0.10	2.0	0.2	0.20	1.0; 1.0	0.4
Purification eluent		Product			
1: Hex / EtOAc (normal phase)		m[Pdt]/mg	n[Pdt]/mmol	yield	TTN
2: water / MeOH (reverse phase)		21.3	0.096	48%	260

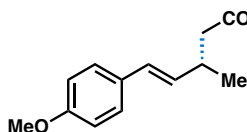


Spectral data is in agreement with literature report<sup>32</sup>. **<sup>1</sup>H NMR** (500 MHz, CDCl<sub>3</sub>) δ 7.14 (d, *J* = 8.5 Hz, 2H), 6.84 (d, *J* = 8.7 Hz, 2H), 4.08 (qd, *J* = 7.2, 1.2 Hz, 2H), 3.79 (s, 3H), 3.24 (h, *J* = 7.1 Hz, 1H), 2.57 (dd, *J* = 14.9, 7.2 Hz, 1H), 2.51 (dd, *J* = 14.9, 8.0 Hz, 1H), 1.28 (d, *J* = 7.0 Hz, 3H), 1.19 (t, *J* = 7.1 Hz, 3H). **<sup>13</sup>C NMR** (126 MHz, CDCl<sub>3</sub>) δ 172.6, 158.2, 138.0, 127.8, 113.9, 60.4, 55.4, 43.4, 35.9, 22.1, 14.3. [ $\alpha$ ]<sub>D</sub><sup>23</sup> = +26.334 ± 0.676° (*c* 0.5, CHCl<sub>3</sub>). **SFC Chiralcel OB-H column** (supercritical CO<sub>2</sub>, 2.5 mL/min, 40 °C), *t<sub>r</sub>* = 5.50 min (minor), 6.24 min (major), 97.9 : 2.1 e.r.

**Ethyl (*E*)-5-(4-methoxyphenyl)-3-methylpent-4-enoate (54d)**

This compound was prepared using **General Procedure D-I** and Workup Procedure D-ii.

<i>E. coli</i> suspension in M9-N (variant: P411-CHF, OD <sub>600</sub> = 32)				GOX solution <sup>a</sup>	<i>D</i> -glucose in M9-N <sup>b</sup>
Addition 1, volume/ mL	Addition 2, volume/mL	[PC]/μM	n_pro/μmol	volume/mL	volume/mL
137.5	70.0	2.59	0.536	12.5	25.0
Alkane substrate (53d) stock in EtOH			Ethyl diazoacetate stock in EtOH		
stock/M	volume/mL	n_1/mmol	stock/M	volume/mL add. 1; add. 2	n_2/mmol
0.10	2.5	0.25	0.20	1.25; 1.25	0.50
Purification eluent		Product			
Hex / (4 : 1 DCM : Et <sub>2</sub> O) (normal phase)		m[Pdt]/mg	n[Pdt]/mmol	yield	TTN
		9.7	0.039	15.6%	70

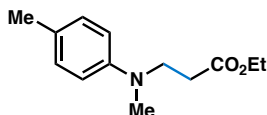


<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.30 – 7.24 (m, 2H), 6.84 (d, *J* = 8.8 Hz, 2H), 6.34 (d, *J* = 15.9 Hz, 1H), 5.99 (dd, *J* = 15.9, 7.6 Hz, 1H), 4.12 (q, *J* = 7.1 Hz, 2H), 3.80 (s, 3H), 2.90 – 2.75 (m, 1H), 2.41 (dd, *J* = 14.7, 7.3 Hz, 1H), 2.34 (dd, *J* = 14.7, 7.3 Hz, 1H), 1.23 (t, *J* = 7.1 Hz, 3H), 1.14 (d, *J* = 6.7 Hz, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 172.6, 159.0, 132.3, 130.4, 128.3, 127.3, 114.0, 60.4, 55.4, 42.1, 34.3, 20.5, 14.5. HRMS (FAB) *m/z*: 248.1417 (M<sup>+</sup>); calc. for C<sub>15</sub>H<sub>20</sub>O<sub>3</sub>: 248.1413. SFC Chiralcel OB-H column (3% *i*-PrOH in supercritical CO<sub>2</sub>, 2.5 mL/min, 40 °C), *t<sub>r</sub>* = 6.62 min (minor), 7.66 min (major), 97.0 : 3.0 e.r.

**Ethyl 3-(methyl(*p*-tolyl)amino)propanoate (57a)**

This compound was prepared using **General Procedure D-I** and Workup Procedure D-i.

<i>E. coli</i> suspension in M9-N (variant: P411-CHF, OD <sub>600</sub> = 29)				GOX solution <sup>a</sup>	<i>D</i> -glucose in M9-N <sup>b</sup>
Addition 1, volume/ mL	Addition 2, volume/mL	[PC]/μM	n_pro/μmol	volume/mL	volume/mL
55.0	28.0	2.13	0.177	5.0	10.0
Alkane substrate (56a) stock in EtOH			Ethyl diazoacetate stock in EtOH		
stock/M	volume/mL	n_1/mmol	stock/M	volume/mL add. 1; add. 2	n_2/mmol
0.50	1.0	0.50	1.00	0.50; 0.50	1.0
Purification eluent		Product			
<b>1:</b> Hex / EtOAc (normal phase)		m[Pdt]/mg	n[Pdt]/mmol	yield	TTN
		91.2	0.412	82%	2330

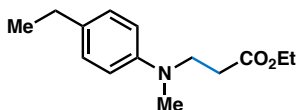


Spectral data is in agreement with literature report<sup>35</sup>. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.05 (d, *J* = 8.3 Hz, 2H), 6.67 (d, *J* = 8.6 Hz, 2H), 4.12 (q, *J* = 7.1 Hz, 2H), 3.64 (t, *J* = 7.2 Hz, 2H), 2.90 (s, 3H), 2.54 (t, *J* = 7.2 Hz, 2H), 2.25 (s, 3H), 1.25 (t, *J* = 7.1 Hz, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 172.6, 146.8, 129.9, 126.3, 113.1, 60.7, 49.1, 38.5, 31.8, 20.4, 14.3.

### Ethyl 3-((4-ethylphenyl)(methyl)amino)propanoate (57b)

This compound was prepared using **General Procedure D-I** and Workup Procedure D-ii.

<i>E. coli</i> suspension in M9-N (variant: P411-CHF, OD <sub>600</sub> = 30)				GOX solution <sup>a</sup>	<i>D</i> -glucose in M9-N <sup>a</sup>
Addition 1, volume/ mL	Addition 2, volume/mL	[PC]/μM	n_pro/μmol	volume/mL	volume/mL
55.0	28.0	2.02	0.168	5.0	10.0
Alkane substrate (56b) stock in EtOH			Ethyl diazoacetate stock in EtOH		
stock/M	volume/mL	n_1/mmol	stock/M	volume/mL add. 1; add. 2	n_2/mmol
0.50	1.0	0.50	1.00	0.50; 0.50	1.0
Purification eluent		Product			
<b>1:</b> Hex / (4 : 1 DCM : Et <sub>2</sub> O) (normal phase)		m[Pdt]/mg	n[Pdt]/mmol	yield	TTN
<b>2:</b> water / MeOH (reverse phase)		88.0	0.374	75%	2230

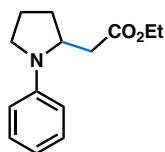


<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.08 (d, *J* = 8.7 Hz, 2H), 6.69 (d, *J* = 8.8 Hz, 2H), 4.13 (q, *J* = 7.1 Hz, 2H), 3.65 (t, *J* = 7.2 Hz, 2H), 2.91 (s, 3H), 2.60 – 2.51 (m, 4H), 1.25 (t, *J* = 7.1 Hz, 3H), 1.20 (t, *J* = 7.6 Hz, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 172.6, 146.9, 132.8, 128.7, 113.0, 60.7, 49.1, 38.5, 31.9, 27.9, 16.1, 14.3. HRMS (ESI-TOF) *m/z*: 236.1673 (*M* + H<sup>+</sup>); calc. for C<sub>14</sub>H<sub>22</sub>NO<sub>2</sub>: 236.1651.

**Ethyl 2-(1-phenylpyrrolidin-2-yl)acetate (57c)**

This compound was prepared using **General Procedure D-I** and Workup Procedure D-ii.

<i>E. coli</i> suspension in M9-N (variant: P411-CHF, OD <sub>600</sub> = 31)				GOX solution <sup>a</sup>	D-glucose in M9-N <sup>b</sup>
Addition 1, volume/ mL	Addition 2, volume/mL	[PC]/μM	n_pro/μmol	volume/mL	volume/mL
55.0	28.0	2.65	0.220	5.0	10.0
Alkane substrate (56c) stock in EtOH			Ethyl diazoacetate stock in EtOH		
stock/M	volume/mL	n_1/mmol	stock/M	volume/mL add. 1; add. 2	n_2/mmol
0.50	1.0	0.50	1.00	0.50; 0.50	1.0
Purification eluent		Product			
<b>1:</b> Hex / (4 : 1 DCM : Et <sub>2</sub> O) (normal phase)		m[Pdt]/mg	n[Pdt]/mmol	yield	TTN
		104.1	0.446	89%	2030

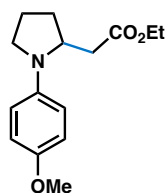


<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.29 – 7.20 (m, 2H), 6.69 (tt, *J* = 7.2, 1.1 Hz, 1H), 6.62 (d, *J* = 7.9 Hz, 2H), 4.21 – 4.14 (m, 3H), 3.48 – 3.37 (m, 1H), 3.23 – 3.14 (m, 1H), 2.79 (dd, *J* = 15.0, 2.9 Hz, 1H), 2.22 (dd, *J* = 15.0, 10.5 Hz, 1H), 2.12 – 1.98 (m, 3H), 1.97 – 1.82 (m, 1H), 1.29 (t, *J* = 7.1 Hz, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 172.2, 146.6, 129.5, 116.0, 112.0, 60.6, 55.5, 48.0, 37.8, 31.1, 23.1, 14.4. HRMS (ESI-TOF) *m/z*: 234.1491 (*M* + *H*<sup>+</sup>); calc. for C<sub>14</sub>H<sub>20</sub>NO<sub>2</sub>: 234.1494. [ $\alpha$ ]<sub>D</sub><sup>23</sup> = +2.056 ± 0.834° (*c* 0.5, CHCl<sub>3</sub>). HPLC Chiralcel OD-H column (6% *i*-PrOH in *n*-hexane, 1.0 mL/min, room temperature), *t*<sub>r</sub> = 6.20 min (minor), 8.58 min (major), 82.8 : 17.2 e.r.

**Ethyl 2-(1-(4-methoxyphenyl)pyrrolidin-2-yl)acetate (57d)**

This compound was prepared using **General Procedure D-I** and Workup Procedure D-i.

<i>E. coli</i> suspension in M9-N (variant: P411-CHF, OD <sub>600</sub> = 31)				GOX solution <sup>a</sup>	D-glucose in M9-N <sup>b</sup>
Addition 1, volume/ mL	Addition 2, volume/mL	[PC]/μM	n_pro/μmol	volume/mL	volume/mL
55.0	28.0	2.17	0.180	5.0	10.0
Alkane substrate (56d) stock in EtOH			Ethyl diazoacetate stock in EtOH		
stock/M	volume/mL	n_1/mmol	stock/M	volume/mL add. 1; add. 2	n_2/mmol
0.50	1.0	0.50	1.0	0.50; 0.50	1.0
Purification eluent		Product			
<b>1:</b> Hex / EtOAc (normal phase) <b>2:</b> water / MeOH (reverse phase)		m[Pdt]/mg	n[Pdt]/mmol	yield	TTN
		102.1	0.388	78%	2150

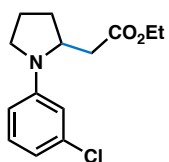


<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.86 (d, *J* = 9.1 Hz, 2H), 6.58 (d, *J* = 9.1 Hz, 2H), 4.16 (q, *J* = 7.2 Hz, 2H), 4.13 – 4.05 (m, 1H), 3.76 (s, 3H), 3.42 – 3.35 (m, 1H), 3.17 – 3.09 (m, 1H), 2.76 (dd, *J* = 14.9, 3.0 Hz, 1H), 2.20 (dd, *J* = 14.9, 10.4 Hz, 1H), 2.13 – 1.96 (m, 3H), 1.92 – 1.82 (m, 1H), 1.28 (t, *J* = 7.1 Hz, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 172.3, 151.1, 141.5, 115.3, 112.9, 60.6, 56.1, 56.0, 48.6, 38.1, 31.1, 23.3, 14.4. HRMS (FAB) *m/z*: 263.1518 (M<sup>+</sup>); calc. for C<sub>15</sub>H<sub>21</sub>NO<sub>3</sub>: 263.1521. [α]<sub>D</sub><sup>23</sup> = +7.310 ± 0.478° (*c* 0.5, CHCl<sub>3</sub>). HPLC Chiralcel OD-H column (6% *i*-PrOH in *n*-hexane, 1.0 mL/min, room temperature), *t*<sub>r</sub> = 7.25 min (minor), 8.12 min (major), 83.7 : 16.3 e.r.

### Ethyl 2-(1-(3-chlorophenyl)pyrrolidin-2-yl)acetate (57e)

This compound was prepared using General Procedure D-I and Workup Procedure D-i.

<i>E. coli</i> suspension in M9-N (variant: P411-CHF, OD <sub>600</sub> = 29)				GOX solution <sup>a</sup>	<i>D</i> -glucose in M9-N <sup>b</sup>
Addition 1, volume/ mL	Addition 2, volume/mL	[PC]/μM	n_pro/μmol	volume/mL	volume/mL
137.5	70.0	3.24	0.673	12.5	25.0
Alkane substrate (56e) stock in EtOH			Ethyl diazoacetate stock in EtOH		
stock/M	volume/mL	n_1/mmol	stock/M	volume/mL add. 1; add. 2	n_2/mmol
0.20	2.5	0.50	0.40	1.25; 1.25	1.0
Purification eluent		Product			
1: Hex / EtOAc (normal phase)		m[Pdt]/mg	n[Pdt]/mmol	yield	TTN
2: water / MeOH (reverse phase)		89.6	0.335	67%	500



<sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>) δ 7.13 (t, *J* = 8.1 Hz, 1H), 6.65 (dd, *J* = 7.9, 1.8 Hz, 1H), 6.57 (t, *J* = 2.2 Hz, 1H), 6.48 (dd, *J* = 8.3, 2.3 Hz, 1H), 4.23 – 4.12 (m, 3H), 3.44 – 3.36 (m, 1H), 3.22 – 3.12 (m, 1H), 2.74 (dd, *J* = 15.1, 3.0 Hz, 1H), 2.23 (dd, *J* = 15.1, 10.4 Hz, 1H), 2.13 – 2.01 (m, 3H), 1.97 – 1.85 (m, 1H), 1.30 (t, *J* = 7.1 Hz, 3H). <sup>13</sup>C NMR (126 MHz, CDCl<sub>3</sub>) δ 171.9, 147.6, 135.3, 130.4, 115.9, 111.9, 110.2, 60.7, 55.5, 48.1, 37.7, 31.1, 23.0, 14.4. HRMS (ESI-TOF) *m/z*: 268.1094 (M + H<sup>+</sup>); calc. for C<sub>14</sub>H<sub>19</sub>NO<sub>2</sub><sup>35</sup>Cl: 268.1104. [α]<sub>D</sub><sup>23</sup> = +7.144 ± 0.875° (*c* 0.5, CHCl<sub>3</sub>). HPLC Chiralcel OD-H column (6% *i*-PrOH in *n*-hexane, 1.0 mL/min, room temperature), *t*<sub>r</sub> = 6.29 min (minor), 6.87 min (major), 90.3 : 9.7 e.r.

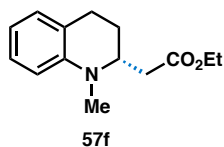
(-)-Ethyl (*R*)-2-(1-methyl-1,2,3,4-tetrahydroquinolin-2-yl)acetate ((-)-**57f**)

This compound was prepared using **General Procedure D-I** and Workup Procedure D-*i*.

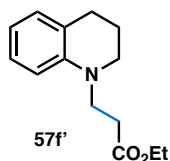
<i>E. coli</i> suspension in M9-N (variant: P411-CHF, OD <sub>600</sub> = 31)				GOX solution <sup>a</sup>	D-glucose in M9-N <sup>b</sup>
Addition 1, volume/ mL	Addition 2, volume/mL	[PC]/μM	n_pro/μmol	volume/mL	volume/mL
55.0	28.0	2.40	0.199	5.0	10.0
Alkane substrate ( <b>56f</b> ) stock in EtOH			Ethyl diazoacetate stock in EtOH		
stock/M	volume/mL	n_1/mmol	stock/M	volume/mL add. 1; add. 2	n_2/mmol
0.50	1.0	0.50	1.00	0.50; 0.50	1.0
Purification eluent		Product			
<b>1:</b> Hex / (4 : 1 DCM : Et <sub>2</sub> O) (normal phase)		m[Pdt]/mg	n[Pdt]/mmol	yield	TTN
<b>2:</b> water / acetonitrile (reverse phase preparative HPLC)		48.9 <sup>c</sup>	0.210	42%	1050 <sup>d</sup>

<sup>c</sup> Regiomer ratio 9 : 1 for **57f** : **57f'** determined by <sup>1</sup>H NMR.

<sup>d</sup> Reported for the sum of regioisomers **57f** and **57f'**.



9 : 1 r.r. (**57f** : **57f'**)



<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.14 – 7.02 (m, 1H, both isomers), 6.99 (d, *J* = 7.3 Hz, 0.9H, isomer **57f**), 6.97 – 6.93 (m, 0.1H, isomer **57f'**), 6.68 – 6.51 (m, 2H, both isomers), 4.15 (q, *J* = 7.1 Hz, 2H, both isomers), 3.86 – 3.78 (m, 0.9H, isomer **57f**), 3.65 – 3.57 (m, 0.2H, isomer **57f'**), 3.32 – 3.25 (m, 0.2H, isomer **57f'**), 2.93 (s, 2.7H, isomer **57f**), 2.91 – 2.79 (m, 0.9H, isomer **57f**), 2.77 – 2.67 (m, 1.1H, both isomers), 2.65 – 2.54 (m, 1.1H, both isomers), 2.39 (dd, *J* = 14.7, 8.6 Hz, 0.9H, isomer **57f**), 2.08 – 1.84 (m, 2H, both isomers), 1.30 – 1.24 (m, 3H, both isomers). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 172.6, 172.3, 144.8, 144.6, 129.5, 129.0, 127.3, 127.3, 122.8, 121.6, 116.2, 116.2, 111.0, 110.7, 60.7, 60.7, 56.1, 49.6, 47.3, 37.8, 36.7, 31.5, 28.2, 25.5, 23.3, 22.3, 14.4. With variant P411-CHF: [α]<sub>D</sub><sup>23</sup> = -2.980 ± 0.898° (*c* 1.0, CHCl<sub>3</sub>). With variant P411-CHF: SFC Chiralcel OB-H column (3% *i*-PrOH supercritical CO<sub>2</sub>, 2.5 mL/min, 40 °C), t<sub>r</sub> = 6.47 min (major), 7.38 min (minor), 73.0 : 27.0 e.r.

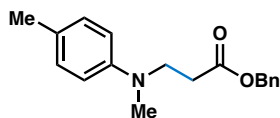
**Benzyl 3-(methyl(*p*-tolyl)amino)propanoate (59a)**

This compound was prepared using **General Procedure D-I** and Workup Procedure D-*i*.

<i>E. coli</i> suspension in M9-N (variant: P411-CHF, OD <sub>600</sub> = 30)				GOX solution <sup>a</sup>	D-glucose in M9-N <sup>b</sup>
Addition 1, volume/ mL	Addition 2, volume/mL	[PC]/μM	n_pro/μmol	volume/mL	volume/mL
55.0	28.0	2.04	0.170	5.0	10.0
Alkane substrate (56a) stock in EtOH			Diazo (58a) stock in EtOH		
stock/M	volume/mL	n_1/mmol	stock/M	volume/mL add. 1; add. 2	n_2/mmol
0.50	1.0	0.50	1.00	0.50; 0.50	1.0
Purification eluent		Product			
1: Hex / EtOAc (normal phase)		m[Pdt]/mg	n[Pdt]/mmol	yield	TTN
		19.0 <sup>c</sup>	0.061 <sup>d</sup>	12% <sup>d</sup>	360 <sup>d</sup>

<sup>c</sup> Isolated with 10% diazo dimer.

<sup>d</sup> Corrected for diazo dimer.



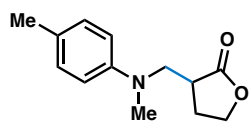
This compound is known in the literature.<sup>36</sup> <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.39 – 7.31 (m, 5H), 7.04 (d, *J* = 8.7 Hz, 2H), 6.66 (d, *J* = 8.8 Hz, 2H), 5.11 (s, 2H), 3.66 (t, *J* = 7.2 Hz, 2H), 2.87 (s, 3H), 2.61 (t, *J* = 7.2 Hz, 2H), 2.25 (s, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 172.4, 146.7, 135.9, 133.9, 129.9, 128.7, 128.4, 126.3, 113.2, 66.5, 49.1, 38.5, 31.8, 20.4. HRMS (ESI-TOF) *m/z*: 284.1635 (*M* + *H*<sup>+</sup>); calc. for C<sub>18</sub>H<sub>22</sub>NO<sub>2</sub>: 284.1651.

**3-((Methyl(*p*-tolyl)amino)methyl)dihydrofuran-2(3*H*)-one (59b)**

This compound was prepared using **General Procedure D-II** and Workup Procedure D-*i*.

<i>E. coli</i> suspension in M9-N (variant: P411-IY (T327I), OD <sub>600</sub> = 55)			GOX solution <sup>a</sup>	D-glucose in M9-N <sup>b</sup>
Volume/ mL	[PC]/μM	n_pro/μmol	volume/mL	volume/mL
88.0	6.77	0.595	0.0	10.0
Alkane substrate (56a) stock in EtOH			Diazo (58b) stock in EtOH	
stock/M	volume/mL	n_1/mmol	stock/M	volume/mL n_2/mmol
1.0	1.0	1.0	0.50	1.0 0.50
Purification eluent		Product		
1: Hex / EtOAc (normal phase)		m[Pdt]/mg	n[Pdt]/mmol	yield
		37.2	0.170	34%
				TTN
				280



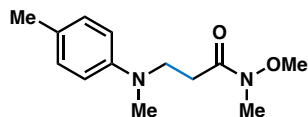


**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.07 (d, *J* = 8.3 Hz, 2H), 6.68 (d, *J* = 8.6 Hz, 2H), 4.35 (td, *J* = 8.8, 2.6 Hz, 1H), 4.15 (ddd, *J* = 10.0, 9.1, 6.6 Hz, 1H), 3.94 (dd, *J* = 15.1, 4.5 Hz, 1H), 3.43 (dd, *J* = 15.1, 8.0 Hz, 1H), 2.96 (s, 3H), 2.91 (dddd, *J* = 10.6, 8.6, 8.0, 4.5 Hz, 1H), 2.34 (dddd, *J* = 12.7, 8.9, 6.6, 2.6 Hz, 1H), 2.26 (s, 3H), 2.11 (dtd, *J* = 12.8, 10.2, 8.5 Hz, 1H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 178.5, 146.8, 130.0, 126.7, 113.1, 66.7, 53.3, 39.6, 38.7, 28.1, 20.4. **HRMS** (FAB) *m/z*: 218.1172 ([*M* + *H*<sup>+</sup>]<sup>+</sup>); calc. for C<sub>13</sub>H<sub>16</sub>NO<sub>2</sub>: 218.1181. With variant **P411-IY T327I**: **HPLC Chiralcel OD-H column** (6% *i*-PrOH in *n*-hexane, 1.0 mL/min, 32 °C), *t<sub>r</sub>* = 22.826 min (major), 24.286 min (minor), 78.0 : 22.0 e.r.

### *N*-Methoxy-*N*-methyl-3-(methyl(*p*-tolyl)amino)propanamide (59c)

This compound was prepared using **General Procedure D-II** and Workup Procedure D-*i*.

<i>E. coli</i> suspension in M9-N (variant: P411-IY (T327I), OD <sub>600</sub> = 55)			GOX solution <sup>a</sup>	<i>D</i> -glucose in M9-N <sup>b</sup>	
Volume/ mL	[PC]/μM	n <sub>pro</sub> /μmol	volume/mL	volume/mL	
88.0	7.00	0.616	0.0	10.0	
Alkane substrate (56a) stock in EtOH		Diazo (58c) stock in EtOH			
stock/M	volume/mL	n <sub>1</sub> /mmol	stock/M	volume/mL	n <sub>2</sub> /mmol
1.0	1.0	1.0	0.50	1.0	0.50
Purification eluent		Product			
1: Hex / EtOAc (normal phase)		m[Pdt]/mg	n[Pdt]/mmol	yield	TTN
		72.3	0.306	61%	500



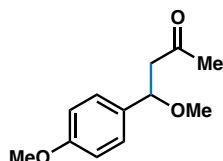
**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.04 (d, *J* = 8.2 Hz, 2H), 6.68 (d, *J* = 8.7 Hz, 2H), 3.70 – 3.64 (m, 2H), 3.62 (s, 3H), 3.17 (s, 3H), 2.92 (s, 3H), 2.66 (t, *J* = 7.2 Hz, 2H), 2.25 (s, 3H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 173.2, 146.8, 129.9, 125.9, 112.8, 61.5, 48.8, 38.6, 32.2, 29.0, 20.4. **HRMS** (ESI-TOF) *m/z*: 237.1616 (*M* + *H*<sup>+</sup>); calc. for C<sub>13</sub>H<sub>21</sub>N<sub>2</sub>O<sub>2</sub>: 237.1603.

#### 4-Methoxy-4-(4-methoxyphenyl)butan-2-one (59d)

This compound was prepared using **General Procedure D-II** and Workup Procedure D-*i*.

<i>E. coli</i> suspension in M9-N (variant: P411-CHF, OD <sub>600</sub> = 60) <sup>c</sup>			GOX solution <sup>a</sup>	<i>D</i> -glucose in M9-N <sup>b</sup>	
Volume/ mL	[PC]/μM	n_pro/μmol	volume/mL	volume/mL	
88.0	8.17	0.719	0.0	10.0	
Alkane substrate (51a) stock in EtOH		Diazo (58d) stock in EtOH			
stock/M	volume/mL	n_1/mmol	stock/M	volume/mL	n_2/mmol
1.0	1.0	1.0	0.50	1.0	0.50
Purification eluent		Product			
<b>1</b> : Hex / EtOAc (normal phase)		m[Pdt]/mg	n[Pdt]/mmol	yield	TTN
		21.8	0.105	21%	150

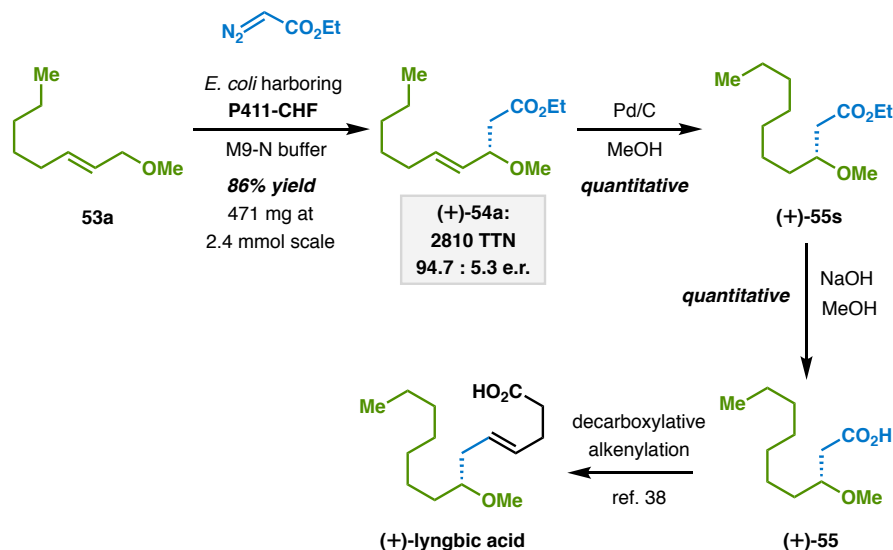
<sup>c</sup> Enzyme was expressed using a modified procedure. After addition of 5-aminolevulinic acid (1.0 mM final concentration) and IPTG (0.5 mM final concentration), the incubator temperature was set to 24 °C, and the culture was allowed to shake for 20 hours at 140 rpm. Product formation was also observed when cultures were expressed following the typical protocol.



<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.24 (d, *J* = 8.7 Hz, 2H), 6.89 (d, *J* = 8.7 Hz, 2H), 4.58 (dd, *J* = 8.8, 4.5 Hz, 1H), 3.81 (s, 3H), 3.17 (s, 3H), 2.97 (dd, *J* = 15.8, 8.8 Hz, 1H), 2.58 (dd, *J* = 15.8, 4.5 Hz, 1H), 2.15 (s, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 206.9, 159.4, 133.0, 127.9, 114.1, 79.3, 56.6, 55.4, 52.0, 31.2. HRMS (ESI-TOF) *m/z*: 231.0999 (*M* + Na<sup>+</sup>); calc. for C<sub>12</sub>H<sub>16</sub>O<sub>3</sub>Na: 231.0997. HPLC Chiralcel OJ-H column (6% *i*-PrOH in *n*-hexane, 1.0 mL/min, 28 °C), *t<sub>r</sub>* = 20.152 min (major), 21.760 min (minor), 71.0 : 29.0 e.r.

### C.9 Syntheses of (+)-lyngbic acid and (+)-cuspareine

Experimental details for the formal synthesis of (*R*)-(+)-lyngbic acid and the total synthesis of (*R*)-(+)-cuspareine can be found below. Enzymatic C–H alkylation is the key stereo-defining step in the routes to both molecules.



**Figure C-12.** Detailed scheme for the formal synthesis of (+)-lyngbic acid.

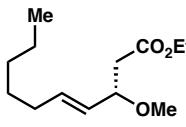
#### (+)-Ethyl (*E*)-3-methoxydec-4-enoate ((+)-54a)

Prepared following **General Procedure D-I** and Workup Procedure D-ii (see **Section C.8**).

<i>E. coli</i> suspension in M9-N (variant: P411-CHF, OD <sub>600</sub> = 29)				GOX solution <sup>a</sup>	<i>D</i> -glucose in M9-N <sup>b</sup>
Addition 1, volume/ mL	Addition 2, volume/mL	[PC]/μM	n_pro/μmol	volume/mL	volume/mL
220.0	112.0	2.21	0.734	20.0	40.0
Alkane substrate (53a) stock in EtOH			Ethyl diazoacetate stock in EtOH		
stock/M	volume/mL	n_1/mmol	stock/M	volume/mL add. 1; add. 2	n_2/mmol
0.60	4.0	2.4	1.2	2.0; 2.0	4.8
Purification eluent		Product			
<b>1:</b> Hex / (4 : 1 DCM : Et <sub>2</sub> O) (normal phase)		m[Pdt]/mg	n[Pdt]/mmol	yield	TTN
<b>2:</b> water / MeOH (reverse phase)		470.6	2.061	86%	2810

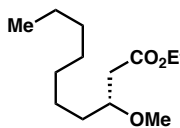
**Notes:** [PC] = protein concentration in original cell suspension, n\_pro = amount of protein in the reaction, n\_1 = amount of alkane substrate in the reaction, n\_2 = total amount of diazo compound in the reaction, add. = addition, m[Pdt] = mass of product isolated, n[Pdt] = amount of product. <sup>a</sup> GOX oxygen depletion system is 14,000 U/mL catalase and 1,000

U/ mL glucose oxidase in 0.1 M potassium phosphate buffer (pH=8.0); the final reaction mixture contains 700 U/mL catalase and 50 U/mL glucose oxidase. <sup>b</sup>*D*-glucose stock solution is 250 mM; final concentration of *D*-glucose in the reaction is 25 mM. These notes apply for all tables in this section.

 <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 5.69 (dt, *J* = 15.4, 6.8 Hz, 1H), 5.28 (ddt, *J* = 15.4, 8.3, 1.5 Hz, 1H), 4.14 (qd, *J* = 7.2, 0.8 Hz, 2H), 3.97 (td, *J* = 8.2, 5.5 Hz, 1H), 3.25 (s, 3H), 2.59 (dd, *J* = 14.9, 8.1 Hz, 1H), 2.42 (dd, *J* = 14.9, 5.5 Hz, 1H), 2.10 – 1.97 (m, 2H), 1.43 – 1.20 (m, 9H), 0.88 (t, *J* = 6.9 Hz, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 171.2, 135.6, 128.7, 79.0, 60.6, 56.2, 41.5, 32.3, 31.5, 28.9, 22.6, 14.4, 14.2. HRMS (EI) *m/z*: 228.1717 (*M*<sup>+</sup>); calc. for C<sub>13</sub>H<sub>24</sub>O<sub>3</sub>: 228.1726. [α]<sub>D</sub><sup>23</sup> = +7.504 ± 0.733° (*c* 0.5, CHCl<sub>3</sub>). GC CycloSil-B column (110 °C), *t*<sub>r</sub> = 53.72 min (minor), 55.35 min (major), 94.7 : 5.3 e.r. Absolute stereochemistry was assigned after elaboration to (+)-**55**.

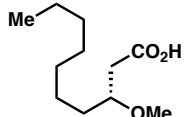
#### (+)-Ethyl 3-methoxydecanoate ((+)-**55s**)

To a solution of (+)-ethyl (*E*)-3-methoxydec-4-enoate ((+)-**54a**, 114.2 mg, 0.5 mmol, 1.0 equiv.) in MeOH (5 mL) was added Pd/C (10% Pd on activated charcoal, 13.3 mg, 2.5 mol%). H<sub>2</sub> was bubbled through the solution for ~30 minutes. The reaction was stirred at room temperature under 1 atm H<sub>2</sub> for 16 hours. The crude product was filtered through a pad of Celite and concentrated under reduced pressure. Purification by silica column chromatography with hexanes / ethyl acetate afforded (+)-**55s** (116.9 mg, 0.5 mmol, quantitative yield). This experiment was conducted by Kai Chen.

 <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.15 (q, *J* = 7.1 Hz, 2H), 3.63 (ddt, *J* = 7.3, 6.3, 5.5 Hz, 1H), 3.35 (s, 3H), 2.52 (dd, *J* = 15.0, 7.3 Hz, 1H), 2.39 (dd, *J* = 15.0, 5.4 Hz, 1H), 1.57 – 1.41 (m, 2H), 1.40 – 1.20 (m, 13H), 0.88 (t, *J* = 7.1 Hz, 3H). <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 172.0, 78.0, 60.5, 57.1, 39.7, 34.1, 31.9, 29.8, 29.4, 25.2, 22.8, 14.4, 14.2. HRMS (FAB) *m/z*: 231.1965 (*M* + H<sup>+</sup>); calc. for C<sub>13</sub>H<sub>27</sub>O<sub>3</sub>: 231.1960. [α]<sub>D</sub><sup>23</sup> = +4.293 ± 0.136° (*c* 1.0, MeOH). Absolute stereochemistry was assigned after derivatization to (+)-**55**.

#### (+)-3-Methoxydecanoic acid ((+)-**55**)

To a solution of (+)-ethyl 3-methoxydecanoate ((+)-**55s**, 46.1 mg, 0.2 mmol, 1.0 equiv.) in MeOH (2 mL) was added NaOH (aq., 15%, 2 mL). The reaction mixture was stirred at room temperature for 1 hour and then slowly acidified with HCl (aq., 1 M) at 0 °C until pH 2–3. Extraction by dichloromethane (15 mL × 3), drying over magnesium sulfate, followed by concentration under reduced pressure afforded product **6** (40.8 mg, 0.2 mmol, quantitative yield) without further purification.

 This compound is known in the literature<sup>38</sup>. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 3.63 (p, *J* = 6.0 Hz, 1H), 3.39 (s, 3H), 2.55 (dd, *J* = 15.5, 6.9 Hz, 1H), 2.50 (dd, *J* = 15.5, 5.3 Hz, 1H), 1.68 – 1.55 (m, 1H), 1.55 – 1.42 (m, 1H),

1.38 – 1.20 (m, 10H), 0.88 (t,  $J = 7.1$  Hz, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  176.2, 77.7, 57.1, 39.1, 33.6, 31.9, 29.7, 29.4, 25.1, 22.8, 14.2.  $[\alpha]_D^{23} = +4.021 \pm 1.649^\circ$  ( $c$  1.0, MeOH). The absolute configuration of (+)-**6** was assigned to be (*R*) by comparing the measured optical rotation value with the literature reported value for (*R*)-(+)-**6** (lit.  $[\alpha]_D^{27} = +3.0^\circ$ ,  $c$  0.67, MeOH)<sup>38</sup>. HRMS (EI)  $m/z$ : 202.1598 ( $\text{M}^+$ ); calc. for  $\text{C}_{11}\text{H}_{22}\text{O}_3$ : 202.1569.

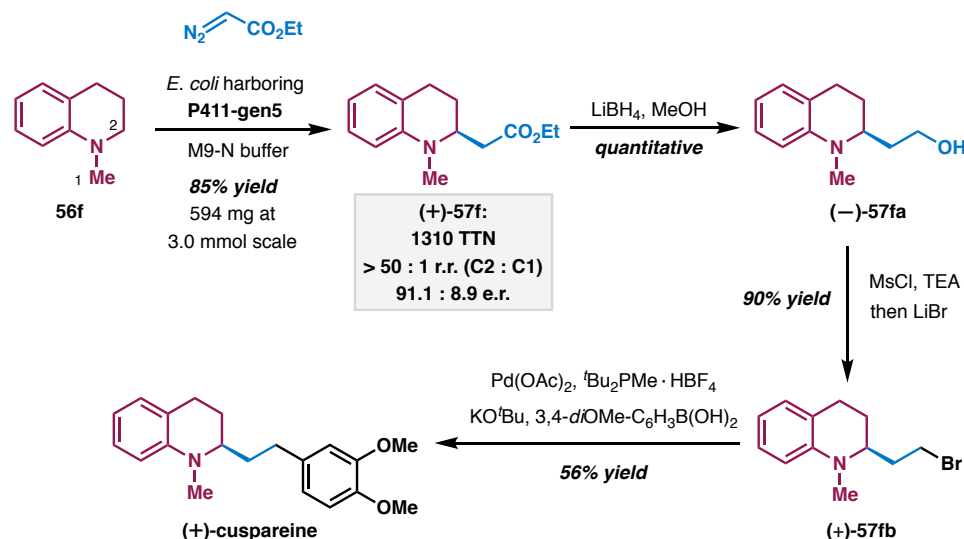


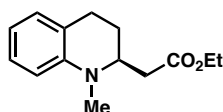
Figure C-13. Detailed scheme for the synthesis of (+)-cuspamine.

(+)-Ethyl (*S*)-2-(1-methyl-1,2,3,4-tetrahydroquinolin-2-yl)acetate ((+)-**57f**)

Prepared following General Procedure D-I and Workup Procedure D-ii (see Section C.8).

<i>E. coli</i> suspension in M9-N (variant: <b>P411-gen5</b> , $\text{OD}_{600} = 31$ )				GOX solution <sup>a</sup>	<i>D</i> -glucose in M9-N <sup>b</sup>
Addition 1, volume/ mL	Addition 2, volume/mL	[PC]/ $\mu\text{M}$	n_pro/ $\mu\text{mol}$	volume/mL	volume/mL
275.0	140.0	4.67	1.938	25.0	50.0
Alkane substrate ( <b>56f</b> ) stock in EtOH			Ethyl diazoacetate stock in EtOH		
stock/M	volume/mL	n_1/mmol	stock/M	volume/mL add. 1; add. 2	n_2/mmol
0.60	5.0	3.0	1.20	2.5; 2.5	6.0
Purification eluent		Product			
1: Hex / (4 : 1 DCM : $\text{Et}_2\text{O}$ )		m[Pdt]/mg	n[Pdt]/mmol	yield	TTN
2: water / MeOH (reverse phase)		594.1 <sup>c</sup>	2.546	85%	1310

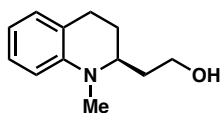
<sup>c</sup> Regiomer ratio > 50 : 1 for **57f** : **57f'** determined by  $^1\text{H}$  NMR.



**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$  7.14 – 7.04 (m, 1H), 6.99 (d,  $J$  = 7.3 Hz, 1H), 6.63 (td,  $J$  = 7.3, 1.1 Hz, 1H), 6.55 (d,  $J$  = 8.2 Hz, 1H), 4.15 (q,  $J$  = 7.1 Hz, 2H), 3.86 – 3.78 (m, 1H), 2.93 (s, 3H), 2.91 – 2.79 (m, 1H), 2.71 (ddd,  $J$  = 16.6, 5.5, 2.8 Hz, 1H), 2.60 (ddd,  $J$  = 14.7, 5.4, 0.8 Hz, 1H), 2.39 (dd,  $J$  = 14.7, 8.6 Hz, 1H), 2.07 – 1.95 (m, 1H), 1.89 (ddt,  $J$  = 13.4, 5.8, 2.9 Hz, 1H), 1.27 (t,  $J$  = 7.1 Hz, 3H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>)  $\delta$  172.3, 144.8, 129.0, 127.3, 121.6, 116.2, 111.0, 60.7, 56.1, 37.8, 36.6, 25.5, 23.3, 14.4. **HRMS** (FAB)  $m/z$ : 233.1405 ( $M^+$ ); calc. for C<sub>14</sub>H<sub>19</sub>NO<sub>2</sub>: 233.1416. With variant **P411-gen5**:  $[\alpha]^{23}_D$  = +9.440  $\pm$  0.292° ( $c$  0.5, CHCl<sub>3</sub>). With variant **P411-gen5**: **SFC Chiralcel OB-H column** (supercritical CO<sub>2</sub>/isopropanol = 97:3, 2.5 mL/min, 40 °C),  $t_r$  = 6.44 min (minor), 7.25 min (major), 91.1 : 8.9 e.r. Absolute stereochemistry was assigned after elaboration to (+)-cuspareine.

**(-)-2-(1-Methyl-1,2,3,4-tetrahydroquinolin-2-yl)ethan-1-ol ((-)-57fa)**

To a solution of (+)-ethyl 2-(1-methyl-1,2,3,4-tetrahydroquinolin-2-yl)acetate ((+)-**57f**, 233.3 mg, 1.0 mmol, 1.0 equiv.) and MeOH (60.8  $\mu$ L, 1.5 mmol, 1.5 equiv.) in dry diethyl ether (20 mL) was added LiBH<sub>4</sub> (2 M in THF, 1.0 mL, 2.0 mmol, 2.0 equiv.) dropwise. The reaction mixture was heated to reflux (~45 °C) for 4 hours (monitored by TLC). Upon completion, the reaction was quenched by ethyl acetate (2 mL) and stirred for 30 min at room temperature, quenched by MeOH (2 mL) and stirred for another 20 min at 0 °C, and finally quenched by NH<sub>4</sub>Cl (sat. aq., 5 mL) at 0 °C. The crude product was then extracted by diethyl ether (30 mL  $\times$  3). The combined organic layer was washed by brine (30 mL), dried over sodium sulfate and concentrated under reduced pressure. Purification by silica column chromatography with hexanes / ethyl acetate afforded **57fa** (192.1 mg, 1.0 mmol, quantitative yield). This experiment was conducted by Kai Chen.

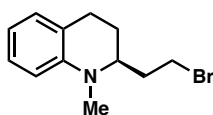


This compound is known in the literature<sup>39</sup>. **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$  7.12 – 7.06 (m, 1H), 7.01 – 6.96 (m, 1H), 6.62 (td,  $J$  = 7.3, 1.2 Hz, 1H), 6.59 (dd,  $J$  = 8.2, 1.1 Hz, 1H), 3.82 – 3.71 (m, 2H), 3.51 – 3.44 (m, 1H), 2.97 (s, 3H), 2.88 – 2.76 (m, 1H), 2.70 (ddd,  $J$  = 16.4, 5.3, 3.4 Hz, 1H), 2.00 – 1.81 (m, 3H), 1.72 – 1.61 (m, 2H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>)  $\delta$  145.4, 129.0, 127.3, 122.2, 116.2, 112.0, 60.6, 56.2, 39.1, 34.6, 24.8, 23.8.  $[\alpha]^{23}_D$  = -28.096  $\pm$  2.416° ( $c$  0.5, CHCl<sub>3</sub>). Absolute stereochemistry was assigned after derivatization to (+)-cuspareine.

**(+)-2-(2-Bromoethyl)-1-methyl-1,2,3,4-tetrahydroquinoline ((+)-57fb)**

To a solution of (-)-2-(1-Methyl-1,2,3,4-tetrahydroquinolin-2-yl)ethan-1-ol ((-)-**57fa**, 114.8 mg, 0.6 mmol, 1.0 equiv.) and triethylamine (133.8  $\mu$ L, 0.96 mmol, 1.6 equiv.) in dry THF (12 mL) at 0 °C was added methanesulfonyl chloride (MsCl, 70  $\mu$ L, 1.5 equiv.). The reaction mixture was warmed to room temperature over 30 min and stirred for additional 30 min. Upon completion, the reaction was quenched by NaHCO<sub>3</sub> (sat. aq., 10 mL) and H<sub>2</sub>O (10 mL). The mesylated product was then extracted by diethyl ether (30 mL

× 3). The combined organic layer was washed by brine (30 mL), dried over sodium sulfate and concentrated under reduced pressure. The mesylated product was then dissolved in DMF (6 mL) and LiBr (259.5 mg, 3.0 mmol, 5.0 equiv.) was added. The reaction was heated to 70 °C and stirred for 2 hours. Upon completion, the reaction was quenched by H<sub>2</sub>O (30 mL). The crude product was then extracted by diethyl ether (30 mL × 3). The combined organic layer was washed by H<sub>2</sub>O (30 mL) and brine (30 mL), dried over sodium sulfate and concentrated under reduced pressure. Purification by silica column chromatography with hexanes / ethyl acetate afforded **57fb** (136.8 mg, 0.54 mmol, 90% yield). This experiment was conducted by Kai Chen.

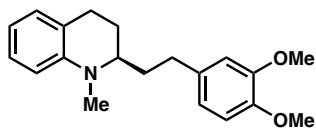


**<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.12 – 7.06 (m, 1H), 7.01 – 6.95 (m, 1H), 6.61 (td, *J* = 7.3, 1.2 Hz, 1H), 6.56 (dd, *J* = 8.2, 1.1 Hz, 1H), 3.58 – 3.47 (m, 2H), 3.43 (ddd, *J* = 10.1, 7.7, 6.3 Hz, 1H), 2.99 (s, 3H), 2.84 – 2.66 (m, 2H), 2.22 – 2.11 (m, 1H), 2.01 – 1.90 (m, 2H), 1.90 – 1.81 (m, 1H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 145.1, 129.0, 127.4, 121.5, 116.0, 111.2, 57.0, 38.7, 34.9, 31.0, 24.3, 23.6. **HRMS** (FAB) *m/z*: 254.0548 (*M* + *H*<sup>+</sup>); calc. for C<sub>12</sub>H<sub>17</sub><sup>79</sup>BrN: 254.0544. [*α*]<sub>D</sub><sup>23</sup> = +33.638 ± 2.022° (*c* 0.5, CHCl<sub>3</sub>). Absolute stereochemistry was assigned after derivatization to (+)-cuspareine.

### (+)-Cuspareine

(+)-Cuspareine was synthesized through Suzuki-Miyaura cross-coupling between an alkyl bromide and an aryl boronic acid. The reaction conditions for this cross-coupling are derived from those described by Fu *et al.*<sup>40</sup>.

To a 50-mL resealable Schlenk tube were added Pd(OAc)<sub>2</sub> (4.04 mg, 0.018 mmol, 15 mol%), di-*tert*-butyl(methyl)phosphonium tetrafluoroborate (*t*Bu<sub>2</sub>PMe·HBF<sub>4</sub>, 8.93 mg, 0.036 mmol, 30 mol%), KOtBu (40.4 mg, 0.36 mmol, 3.0 equiv.) and *t*Amyl-OH (0.6 mL). The tube was charged with Ar and sealed. The mixture was stirred at 60 °C for 20 min until the color of mixture turned pale yellow. After the mixture was cooled to room temperature, (3,4-dimethoxyphenyl)boronic acid (54.6 mg, 0.30 mmol, 2.5 equiv.) and a solution of (+)-2-(2-bromoethyl)-1-methyl-1,2,3,4-tetrahydroquinoline ((+)-**8fb**, 30.5 mg, 0.12 mmol, 1.0 equiv.) in *t*Amyl-OH (0.6 mL) were added to the tube. Another portion of *t*Amyl-OH (0.8 mL) was used to wash the solution residue of (+)-**8fb** and then transferred to the tube. The tube was charged with Ar and sealed again. After the reaction mixture was stirred at 60 °C for 15 hours, it was then cooled to room temperature, diluted with diethyl ether (6 mL), filtrated through a pad of Celite, washed by diethyl ether (30 mL) and concentrated under reduced pressure. Purification by silica column chromatography with hexanes / ethyl acetate afforded (+)-cuspareine (20.8 mg, 0.067 mmol, 56% yield). This experiment was conducted by Kai Chen.



Spectral data is in agreement with literature report.<sup>41</sup> **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.09 (td, *J* = 7.6, 1.4 Hz, 1H), 6.99 (d, *J* = 7.0 Hz, 1H), 6.80 (d, *J* = 8.0 Hz, 1H), 6.75 – 6.70 (m, 2H), 6.60 (td, *J* = 7.3, 1.1 Hz, 1H), 6.54 (d, *J* = 8.2 Hz, 1H), 3.88 (s, 3H), 3.86 (s, 3H), 3.33 – 3.26 (m, 1H), 2.92 (s, 3H), 2.91 – 2.80 (m, 1H), 2.74 – 2.63 (m, 2H), 2.54 (ddd, *J* = 13.9, 10.1, 6.4 Hz, 1H), 2.01 – 1.87 (m, 3H), 1.74 (dddd, *J* = 13.6, 10.1, 8.8, 5.4 Hz, 1H). **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 149.0, 147.3, 145.4, 134.8, 128.8, 127.2, 121.8, 120.2, 115.5, 111.7, 111.4, 110.7, 58.5, 56.1, 56.0, 38.2, 33.2, 32.1, 24.5, 23.7. **[α]<sub>D</sub><sup>23</sup>** = +23.404 ± 0.723° (*c* 0.5, CHCl<sub>3</sub>). The absolute configuration of (+)-cuspareine was assigned to be (*R*) by comparing the measured optical rotation value with the literature reported value for (*R*)-(+)-cuspareine (lit. **[α]<sub>D</sub><sup>27</sup>** = +23.516°, *c* 0.8, CHCl<sub>3</sub>).<sup>41</sup>



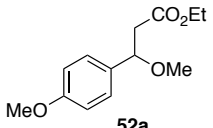
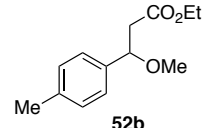
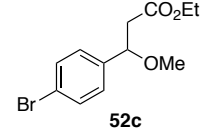
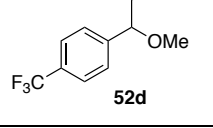
## C.10 Determination of enantioselectivity

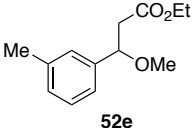
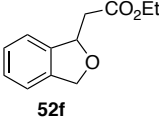
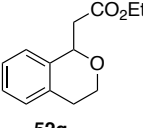
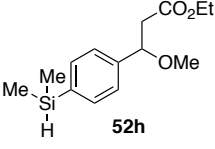
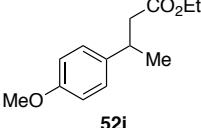
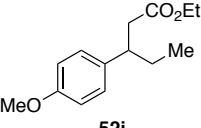
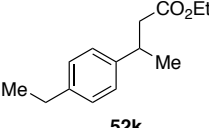
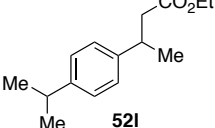
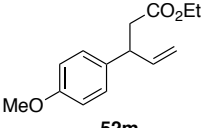
The absolute configuration of P411-CHF synthesized **52i** was assigned to be (*S*) by comparing the chiral HPLC separation of *rac*-**52i** and P411-CHF synthesized **52i** with that reported in the literature<sup>42</sup>. The absolute configuration was further confirmed by comparing the optical rotation values for P411-CHF synthesized **52i** and (*S*)-(+)-**52i** as reported in the literature<sup>42</sup>. The absolute configurations of **52a–52h** and **52j–52m** were inferred by analogy, assuming the selectivity with P411-CHF remains the same in the enzymatic syntheses of those compounds.

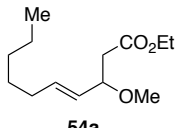
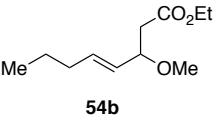
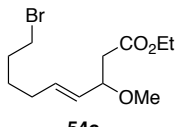
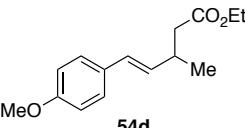
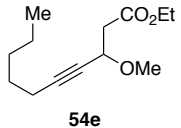
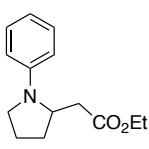
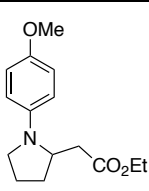
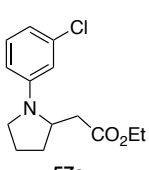
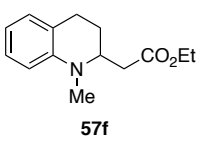
The absolute configuration of P411-CHF synthesized (+)-**54a** was determined to be (*S*) after elaboration to (+)-**55** (see **Section C.9** for more details). The absolute configurations of **54b–54e** were inferred by analogy, assuming the selectivity with P411-CHF remains the same in the enzymatic syntheses of those compounds.

The absolute configuration of P411-gen5 synthesized (+)-**57f** was determined to be (*S*) after elaboration to (+)-cuspamine (see **Section C.9** for more details). As P411-CHF shows the opposite selectivity for the synthesis of (–)-**57f**, P411-CHF produces (*R*)-**57f** in excess.

Enantioselectivity of enzymatic C–H alkylation products were determined by chiral HPLC, chiral SFC, or chiral GC analysis. Conditions are shown below. Representative traces and further information for compounds **57c**, **57d**, and **57e**, can be found in the Supplementary Information of the published paper.

Compound	Analytical conditions for separating enantiomers
 <b>52a</b>	SFC (Chiralpak AD-H column): 3% isopropanol in supercritical CO <sub>2</sub> , 2.5 mL/min, 40 °C, 235 nm. t <sub>R</sub> : 9.1 min, 10.5 min
 <b>52b</b>	SFC (Chiralpak AD-H column): 2% isopropanol in supercritical CO <sub>2</sub> , 2.5 mL/min, 40 °C, 210 nm t <sub>R</sub> : 5.7 min, 7.3 min
 <b>52c</b>	SFC (Chiralpak AD-H column): 3% isopropanol in supercritical CO <sub>2</sub> , 2.5 mL/min, 40 °C, 210 nm t <sub>R</sub> : 8.1 min, 9.3 min
 <b>52d</b>	GC (Chiraldex G-TA column): 110 °C t <sub>R</sub> : 36.4 min, 37.6 min

 <p><b>52e</b></p>	<p>HPLC (Chiralpak IC column): 4% isopropanol in hexane, 1.5 mL/min, 23 °C, 210 nm</p> <p><math>t_R</math>: 5.7 min, 6.8 min</p>
 <p><b>52f</b></p>	<p>SFC (Chiralpak AD-H column): 3% isopropanol in supercritical CO<sub>2</sub>, 2.5 mL/min, 40 °C, 210 nm</p> <p><math>t_R</math>: 8.5 min, 10.4 min</p>
 <p><b>52g</b></p>	<p>SFC (Chiralcel OJ-H column): 0.2% isopropanol in supercritical CO<sub>2</sub>, 2.5 mL/min, 40 °C, 210 nm</p> <p><math>t_R</math>: 5.3 min, 6.0 min</p>
 <p><b>52h</b></p>	<p>SFC (Chiralcel OD-H column): 2% isopropanol in supercritical CO<sub>2</sub>, 2.5 mL/min, 40 °C, 210 nm</p> <p><math>t_R</math>: 3.3 min, 4.0 min</p>
 <p><b>52i</b></p>	<p>SFC (Chiralcel OB-H column): supercritical CO<sub>2</sub>, 2.5 mL/min, 40 °C, 210 nm</p> <p><math>t_R</math>: 5.5 min, 6.4 min</p>
 <p><b>52j</b></p>	<p>SFC (Chiralcel OB-H column): 1% isopropanol in supercritical CO<sub>2</sub>, 2.5 mL/min, 40 °C, 210 nm</p> <p><math>t_R</math>: 3.8 min, 4.6 min</p>
 <p><b>52k</b></p>	<p>HPLC (Chiralcel OD-H column): 3% isopropanol in hexane, 1.0 mL/min, 32 °C, 210 nm</p> <p><math>t_R</math>: 6.2 min, 6.6 min</p>
 <p><b>52l</b></p>	<p>SFC (Chiralpak IC column): 0.5% isopropanol in supercritical CO<sub>2</sub>, 2.5 mL/min, 40 °C, 210 nm</p> <p><math>t_R</math>: 7.9 min, 8.7 min</p>
 <p><b>52m</b></p>	<p>SFC (Chiralpak AD-H column): 1% isopropanol in supercritical CO<sub>2</sub>, 2.5 mL/min, 40 °C, 210 nm</p> <p><math>t_R</math>: 6.7 min, 7.5 min</p>

 <p><b>54a</b></p>	<p>GC (CycloSil-B column): 110 °C  <math>t_R</math>: 53.6 min, 55.4 min</p>
 <p><b>54b</b></p>	<p>GC (CycloSil-B column): 90 °C  <math>t_R</math>: 45.6 min, 48.3 min</p>
 <p><b>54c</b></p>	<p>GC (CycloSil-B column): 130 °C  <math>t_R</math>: 108.6 min, 110.6 min</p>
 <p><b>54d</b></p>	<p>SFC (Chiralcel OB-H column): 3% isopropanol in supercritical CO<sub>2</sub>, 2.5 mL/min, 40 °C, 254 nm  <math>t_R</math>: 6.6 min, 7.6 min</p>
 <p><b>54e</b></p>	<p>GC (CycloSil-B column): 110 °C for 50 min., ramp 1 °C per minute for 10 min. (to 120 °C), 120 °C for 10 min.  <math>t_R</math>: 64.1 min, 65.2 min</p>
 <p><b>57c</b></p>	<p>HPLC (Chiralcel OD-H column): 6% isopropanol in hexane, 1.0 mL/min, room temp., 235 nm  <math>t_R</math>: 6.2 min, 8.6 min</p>
 <p><b>57d</b></p>	<p>HPLC (Chiralcel OD-H column): 6% isopropanol in hexane, 1.0 mL/min, room temp., 235 nm  <math>t_R</math>: 7.2 min, 8.1 min</p>
 <p><b>57e</b></p>	<p>HPLC (Chiralcel OD-H column): 6% isopropanol in hexane, 1.0 mL/min, room temp., 235 nm  <math>t_R</math>: 6.3 min, 6.9 min</p>
 <p><b>57f</b></p>	<p>SFC (Chiralcel OB-H column): 3% isopropanol in supercritical CO<sub>2</sub>, 2.5 mL/min, 40 °C, 235 nm  <math>t_R</math>: 6.4 min, 7.3 min</p>

### C.11 $^1\text{H}$ and $^{13}\text{C}$ NMR

$^1\text{H}$  and  $^{13}\text{C}$  NMR spectra of new compounds can be found in the Supplementary Information of the published paper.

### C.12 References and Notes

1. Kan, S. B. J.; Lewis, R. D.; Chen, K.; Arnold, F. H. Directed evolution of cytochrome *c* for carbon–silicon bond formation: Bringing silicon to life. *Science* **2016**, *354*, 1048–1051.
2. Knight, A. M.; Kan, S. B. J.; Lewis, R. D.; Brandenburg, O. F.; Chen, K.; Arnold, F. H. Diverse engineered heme proteins enable stereodivergent cyclopropanation of unactivated alkenes. *ACS Cent. Sci.* **2018**, *4*, 372–377.
3. Hernandez, K. E.; Renata, H.; Lewis, R. D.; Kan, S. B. J.; Zhang, C.; Forte, J.; Rozzell, D.; McIntosh, J. A.; Arnold, F. H. Highly stereoselective biocatalytic synthesis of key cyclopropane intermediate to ticagrelor. *ACS Catal.* **2016**, *6*, 7810–7813.
4. Teh, A.-H.; Saito, J. A.; Baharuddin, A.; Tuckerman, J. R.; Newhouse, J. S.; Kanbe, M.; Newhouse, E. I.; Rahim, R. A.; Favier, F.; Didierjean, C.; Sousa, E. H.S.; Stott, M. B.; Dunfield, P. F.; Gonzalez, G.; Gilles-Gonzalez, M.-A.; Najimudin, N.; Alam, M. Hell's Gate globin I: An acid and thermostable bacterial hemoglobin resembling mammalian neuroglobin. *FEBS Lett.* **2011**, *585*, 3250–3258.
5. Prier, C. K.; Zhang, R. K.; Buller, A. R.; Brinkmann-Chen, S.; Arnold, F. H. Enantioselective, intermolecular benzylic C–H amination catalysed by an engineered iron-haem enzyme. *Nat. Chem.* **2017**, *9*, 629–634.
6. Farwell, C. C.; Zhang, R. K.; McIntosh, J. A.; Hyster, T. K.; Arnold, F. H. Enantioselective enzyme-catalyzed aziridination enabled by active-site evolution of a cytochrome P450. *ACS Cent. Sci.* **2015**, *1*, 89–93.
7. (a) Brandenburg, O. F.; Prier, C. K.; Chen, K.; Knight, A. M.; Wu, Z.; Arnold, F. H. Stereoselective enzymatic synthesis of heteroatom-substituted cyclopropanes. *ACS Catal.* **2018**, *8*, 2629–2634. (b) **This reference provides information about cytochrome P411:** Coelho, P. S.; Wang, Z. J.; Ener, M. E.; Baril, S. A.; Kannan, A.; Arnold, F. H. A serine-substituted P450 catalyzes highly efficient carbene transfer to olefins in vivo. *Nat. Chem. Biol.* **2013**, *9*, 485–487.
8. Kille, S.; Acevedo-Rocha, C. G.; Parra, L. P.; Zhang, Z.-G.; Opperman, D. J.; Reetz, M. T.; Acevedo, J. P. Reducing codon redundancy and screening effort of

- combinatorial protein libraries created by saturation mutagenesis. *ACS Synth. Biol.* **2013**, *2*, 83–92.
9. Hyster, T. K.; Farwell, C. C.; Buller, A. R.; McIntosh, J. A.; Arnold, F. H. Enzyme-controlled nitrogen-atom transfer enables regiodivergent C–H amination. *J. Am. Chem. Soc.* **2014**, *136*, 15505–15508.
  10. Tang, Y.; Chen, Q.; Liu, X.; Wang, G.; Lin, L.; Feng, X. Direct synthesis of chiral allenolates from the asymmetric C–H insertion of  $\alpha$ -diazoesters into terminal alkynes. *Angew. Chem. Int. Ed.* **2015**, *54*, 9512–9516.
  11. Schwartz, B. D.; Denton, J. R.; Lian, Y.; Davies, H. M. L.; Williams, C. M. Asymmetric [4 + 3] cycloadditions between vinylcarbenoids and dienes: application to the total synthesis of the natural product (–)-5-*epi*-Vibsanin E. *J. Am. Chem. Soc.* **2009**, *131*, 8329–8332.
  12. Yin, G.; Wu, Y.; Liu, G. Scope and mechanism of allylic C–H amination of terminal alkenes by the palladium/PhI(OPiv)<sub>2</sub> catalyst system: insights into the effect of naphthoquinone. *J. Am. Chem. Soc.* **2010**, *132*, 11978–11987.
  13. Bakos, M.; Gyömöre, A.; Domján, A.; Soós, T. Auto-tandem catalysis with frustrated lewis pairs for reductive etherification of aldehydes and ketones. *Angew. Chem. Int. Ed.* **2017**, *56*, 5217–5221.
  14. Hattori, K.; Sajiki, H.; Hirota, K. Chemoselective control of hydrogenation among aromatic carbonyl and benzyl alcohol derivatives using Pd/C(en) catalyst. *Tetrahedron* **2001**, *57*, 4817–4824.
  15. Mitsudome, T.; Mizumoto, K.; Mizugaki, T.; Jitsukawa, K.; Kaneda, K. Wacker-type oxidation of internal olefins using a PdCl<sub>2</sub>/N,N-dimethylacetamide catalyst system under copper-free reaction conditions. *Angew. Chem. Int. Ed.* **2010**, *49*, 1238–1240.
  16. Park, S. R.; Kim, C.; Kim, D.-G.; Thrimurtulu, N.; Yeom, H.-S.; Jun, J.; Shin, S., Rhee, Y. H. Entry to  $\beta$ -alkoxyacrylates via gold-catalyzed intermolecular coupling of alkynoates and allylic ethers. *Org. Lett.* **2013**, *15*, 1166–1169.
  17. Harms, A. E.; Stille, J. R.; Taylor, S. K. Ring formation through intramolecular S<sub>N</sub>' displacement of an allylic methoxy substituent. *Organometallics* **1994**, *13*, 1456–1464.
  18. Colomer, I.; Barcelos, R. C.; Christensen, K. E.; Donohoe, T. J. Orthogonally protected 1,2-diols from electron-rich alkenes using metal-free olefin *syn*-dihydroxylation. *Org. Lett.* **2016**, *18*, 5880–5883.

19. Wender, P. A.; Deschamps, N. M.; Williams, T. J. Intermolecular dienyl Pauson–Khand reaction. *Angew. Chem. Int. Ed.* **2004**, *43*, 3076–3079.
20. Huang, L.; Niu, T.; Wu, J.; Zhang, Y. Copper-catalyzed oxidative cross-coupling of *N,N*-dimethylanilines with heteroarenes under molecular oxygen. *J. Org. Chem.* **2011**, *76*, 1759–1766.
21. Sattely, E. S.; Meek, S. J.; Malcolmson, S. J.; Schrock, R. R.; Hoveyda, A. H. Design and stereoselective preparation of a new class of chiral olefin metathesis catalysts and application to enantioselective synthesis of quebrachamine: Catalyst development inspired by natural product synthesis. *J. Am. Chem. Soc.* **2009**, *131*, 943–953.
22. Hari, D. P.; Waser, J. Enantioselective copper-catalyzed oxy-alkynylation of diazo compounds. *J. Am. Chem. Soc.* **2017**, *139*, 8420–8423.
23. Zhang, L.; Sun, B.; Liu, Q.; Mo, F. Addition of diazo compounds *ipso*-C–H bond to carbon disulfide: Synthesis of 1,2,3-thiadiazoles under mild conditions. *J. Org. Chem.* **2018**, *83*, 4275–4278.
24. Huang, L.; Wulff, W. D. Catalytic asymmetric synthesis of trisubstituted aziridines. *J. Am. Chem. Soc.* **2011**, *133*, 8892–8895.
25. Dakarapu, U. S.; Bokka, A.; Asgari, P.; Trog, G.; Hua, Y.; Nguyen, H. H.; Rahman, N.; Jeon, J. Lewis base activation of silyl acetals: iridium-catalyzed reductive Horner–Wadsworth–Emmons olefination. *Org. Lett.* **2015**, *17*, 5792–5795.
26. TenBrink, R. E.; Bergh, C. L.; Duncan, J. N.; Harris, D. W.; Huff, R. M.; Lahti, R. A.; Lawson, C. F.; Lutzke, B. S.; Martin, I. J.; Rees, S. A.; Schlachter, S. K.; Sih, J. C.; Smith, M. W. (*S*)-(-)-4-[4-[2-(Isochroman-1-yl)ethyl]piperazin-1-yl]benzenesulfonamide, a selective dopamine D4 antagonist. *J. Med. Chem.* **1996**, *39*, 2435–2437.
27. Su, B.; Hartwig, J. F. Ir-catalyzed enantioselective, intramolecular silylation of methyl C–H bonds. *J. Am. Chem. Soc.* **2017**, *139*, 12137–12140.
28. Lu, W.-J.; Chen, Y.-W.; Hou, X.-L. Highly enantioselective iridium-catalyzed hydrogenation of trisubstituted olefins,  $\alpha,\beta$ -unsaturated ketones and imines with chiral benzylic substituted *P,N*-ligands. *Adv. Synth. Catal.* **2010**, *352*, 103–107.
29. Huo, X.; Yang, G.; Liu, D.; Liu, Y.; Gridnev, I. D.; Zhang, W. Palladium-catalyzed allylic alkylation of simple ketones with allylic alcohols and its mechanistic study. *Angew. Chem. Int. Ed.* **2014**, *53*, 6776–6780.

- 
30. Koszelewski, D.; Brodzka, A.; Źądło, A.; Paprocki, D.; Trzepizur, D.; Zysk, M.; Ostaszewski, R. Dynamic kinetic resolution of 3-aryl-4-pentenoic acids. *ACS Catal.* **2016**, *6*, 3287–3292.
31. Candish, L.; Lupton, D. W. N-heterocyclic carbene cascade catalysis: Dual Brønsted/Lewis base rearrangement of cyclopropyl enol esters to dihydropyranones. *Chem. Sci.* **2012**, *3*, 380–383.
32. Shevlin, M.; Friedfeld, M. R.; Sheng, H.; Pierson, N. A.; Hoyt, J. M.; Campeau, L.-C.; Chirik, P. J. Nickel-catalyzed asymmetric alkene hydrogenation of  $\alpha,\beta$ -unsaturated esters: high-throughput experimentation-enabled reaction discovery, optimization, and mechanistic elucidation. *J. Am. Chem. Soc.* **2016**, *138*, 3562–3569.
33. Katritzky, A. R.; Strah, S.; Belyakov, S. A. The preparation of functionalized amines and amides using benzotriazole derivatives and organozinc reagents. *Tetrahedron* **1998**, *54*, 7167–7178.
34. Yadav, J. S.; Reddy, B. V. S.; Hashim, S. R. A new and efficient synthesis of 2,2-disubstituted-3,4-dihydro-2H-1-benzopyrans. *J. Chem. Soc., Perkin Trans. 1* **2000**, *0*, 3082–3084.
35. Shao, Z.; Fu, S.; Wei, M.; Zhou, S.; Liu, Q. Mild and selective cobalt-catalyzed chemodivergent transfer hydrogenation of nitriles, *Angew. Chem. Int. Ed.* **2016**, *55*, 14653–14657.
36. Solé, D.; Serrano, O. Palladium-Catalyzed intramolecular nucleophilic substitution at the alkoxycarbonyl group. *Angew. Chem. Int. Ed.* **2007**, *46*, 7270–7272.
37. Edwards, J. T.; Merchant, R. R.; McClymont, K. S.; Knouse, K. W.; Qin, T.; Malins, L. R.; Vokits, B.; Shaw, S. A.; Bao, D.-H.; Wei, F.-L.; Zhou, T.; Eastgate, M. D.; Baran, P. S. Decarboxylative alkenylation. *Nature* **2017**, *545*, 213–218.
38. Tsai, C.-C.; Chuang, W.-T.; Tsai, Y.-F.; Li, J.-T.; Wu, Y.-F.; Liao, C.-C. Intra- and intermolecular hydrogen bonds enhance the fluoride-responsiveness of functionalized glycolipid-based gelators. *J. Mater. Chem. B* **2013**, *1*, 819–827.
39. Melzig, L.; Dennenwaldt, T.; Gavryushin, A.; Knochel, P. Direct aminoalkylation of arenes, heteroarenes, and alkenes via Ni-catalyzed Negishi cross-coupling reactions. *J. Org. Chem.* **2011**, *76*, 8891–8906.
40. Kirchhoff, J. H.; Netherton, M. R.; Hills, I. D.; Fu, G. C. Boronic acids: new coupling partners in room-temperature Suzuki reactions of alkyl bromides. Crystallographic characterization of an oxidative-addition adduct generated under remarkably mild conditions. *J. Am. Chem. Soc.* **2002**, *124*, 13662–13663.

- 
41. Wang, Y.; Liu, Y.; Zhang, D.; Wei, H.; Shi, M.; Wang, F. Enantioselective Rhodium-catalyzed dearomative arylation or alkenylation of quinolinium salts. *Angew. Chem. Int. Ed.* **2016**, *55*, 3776–3780.
  42. Guo, S.; Zhou, J. *N,N*-Dimethylformamide as hydride source in nickel-catalyzed asymmetric hydrogenation of  $\alpha,\beta$ -unsaturated esters. *Org. Lett.* **2016**, *18*, 5344–5347.